

Essays in Development and Behavioral Economics

by

Yi Han

B. A. in Economics, Renmin University of China, 2012

M. A. in Economics, Renmin University of China, 2014

Submitted to the Graduate Faculty of
the Dietrich School of Arts and Sciences in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH
DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Yi Han

It was defended on

April 3rd, 2020

and approved by

Daniel Berkowitz, Department of Economics. University of Pittsburgh

Thomas Rawski, Department of Economics. University of Pittsburgh

Yogita Shamdasani, Department of Economics. University of Pittsburgh

Jason Cook, Department of Economics. University of Pittsburgh

George Loewenstein, Social and Decision Sciences. Carnegie Mellon University

Dissertation Director: Daniel Berkowitz, Department of Economics. University of
Pittsburgh

Copyright © by Yi Han
2020

Essays in Development and Behavioral Economics

Yi Han, PhD

University of Pittsburgh, 2020

Along with the reduction in transportation costs in the last two centuries, institutional trade barriers have become increasingly important obstacles to further market integration. In Chapter 1, I examine the impact of a policy reform in China that removed inter-regional administrative trade barriers by incorporating counties into prefectures with a larger market. Using a difference-in-differences approach, I compare incorporated counties, both before and after the reform, to two novel control groups: counties that applied for incorporation but failed and counties that were incorporated several years later. I find that the reform immediately and persistently increased the economic growth of incorporated counties. Several sources of evidence suggest that treated counties experienced relatively rapid growth because they became more integrated into the domestic market. In Chapter 2, using data from the one-child policy in China (OCP), Yiming Liu and I provide first field evidence for responsibility-shifting through delegation. We compare the impact of the OCP on parents who experienced it during Phase I when local governments were the enforcer, versus Phase II when the enforcement was delegated to civilians by incentivizing them to report neighbors' violations. We find, consistent with responsibility-shifting, exposure to the OCP in Phase I reduces people's trust in local governments, but exposure to it in Phase II only reduces people's trust in neighbors, not their trust in local governments. In chapter 3, George Loewenstein, Yiming Liu and I designed an online experiment to investigate correspondence bias - when drawing inferences about a person's enduring characteristics from her actions, people tend to overly emphasize the role of the person's enduring characteristics and underestimate the influence of transient situational factors. We build a simple model to formalize this bias and test the predictions of the model. We find evidence of the existence of correspondence bias. Moreover, we show experiencing the games by oneself instead of observing it reduces the bias, and providing counterfactual information on how the benign-game

(malign-game) player behaves in the malign-game (benign-game) eliminates it.

Keywords: Market integration, Local protectionism, Delegation, Responsibility-shifting, Correspondence bias.

Table of Contents

1.0 Administrative Barriers, Market Integration and Economic Growth: Evidence from China	1
1.1 Introduction	1
1.2 Background	7
1.2.1 Local Protectionism	7
1.2.2 Institutional Background	8
1.2.3 Two Special Features	10
1.3 Data	13
1.4 Empirical Strategy	14
1.5 Estimates of the Effect of the Reform on Economic Growth	18
1.5.1 DID Comparing Incorporated Counties to Applied-but-failed Counties	18
1.5.2 DID Using Variation in the Timing of Incorporation	19
1.5.3 Robustness	19
1.6 Mechanism	20
1.7 Conclusion	24
1.8 Figures and Tables	27
2.0 Responsibility-Shifting through Delegation: Evidence from China's One-Child Policy	48
2.1 Introduction	48
2.2 Background of the One-Child Policy	53
2.3 Data	57
2.3.1 The Measurement of Trust	57
2.3.2 The Individual-Level OCP Exposure	57
2.4 Empirical Strategy	60
2.4.1 Identification Strategy	60

2.4.2 Empirical Specification	62
2.5 Results	63
2.5.1 The Effect of OCP Exposure between 1991 and 2015 on Trust	63
2.5.2 The Effect of OCP Exposure in the 1980s on Trust	64
2.5.3 Mechanism	66
2.5.4 Robustness Checks	66
2.6 Discussions	67
2.6.1 Dare Not To Report Mistrust	68
2.6.2 Trust in Central Government	68
2.6.3 Better Performance in the Second Phase	70
2.7 Conclusion	70
2.8 Tables	73
3.0 Correspondence Bias	82
3.1 Introduction	82
3.2 Model	88
3.3 Design	91
3.3.1 First Stage	92
3.3.2 Second Stage	92
3.3.3 Third Stage	93
3.3.4 Three Confounders	95
3.3.5 Treatments	96
3.3.6 Predictions	97
3.4 Results	99
3.5 Conclusion	104
3.6 Figures	106
3.7 Tables	111
4.0 Appendices	115
4.1 Appendix for Chapter 1	115
4.2 Appendix for Chapter 2	120
4.2.1 Fertility Penalty Data	120

4.2.2 The One-tailed T Test for Sex Ratios	122
5.0 References	133

List of Tables

1.1	Summary Statistics (Baseline Year)	40
1.2	Estimated Effects of the Reform on Economic Growth	41
1.3	Estimated Effects of the Reform on Economic Growth: Use Time Variation . .	42
1.4	Overall Effect of the Reform on Prefectures's Economic Growth	43
1.5	Mechanism: Geographical Concentration	44
1.6	Mechanism: Inter-sector Reallocation	45
1.7	Mechanism: Firms' Entry	46
1.8	Mechanism: Firms' Exit	47
2.1	Summary Statistics	73
2.2	Fertility Patterns in China	74
2.3	Factors that Predict Gender of the First Child	75
2.4	Estimates of OCP Exposure on Trust in Phase II (1991-2010)	76
2.5	Estimates of OCP Exposure on Trust in Phase I (1979-1985)	77
2.6	Heterogeneous Effect of OCP Exposure on Trust in Phase II	78
2.7	Heterogeneous Effect of OCP Exposure on Trust in Phase I	79
2.8	Estimates of OCP Exposure on Trust in Central Government	80
2.9	Correlation between Fertility Penalty and Government Performance	81
3.1	The Benign and Malign Games	111
3.2	The Multiple Price List	112
3.3	Benign Premium across Treatments	112
3.4	Summary Statistics	113
3.5	Benign Premiums in Treatment 2	114
4.1	Factors that Predict Timing of Incorporations	117
4.2	Estimated Effects of the Reform on Economic Growth: Approach I (Without Sample of Direct-administered Municipalities of China)	118

4.3	Estimated Effects of the Reform on Economic Growth: Approach II (Without Sample of Direct-administered Municipalities of China)	119
4.4	Sex Ratio at First Birth (Males/Females) in Urban China in 2000	123
4.5	Estimates of OCP Exposure on Trust in Phase II (1991-2010): Excluding High Sex Ratio Provinces	124
4.6	Estimates of OCP Exposure on Trust in Phase II	125
4.7	Estimates of OCP Exposure on Trust in Phase I	126
4.8	Estimates of OCP Exposure on Trust in Phase II	127
4.9	Estimates of OCP Exposure on Trust in Phase II	128
4.10	Estimates of OCP Exposure on Trust in Phase I	129
4.11	Estimates of OCP Exposure on Trust in Phase I	130
4.12	Distribution of Responses: Trust in Local Governments	131
4.13	Correlation between Fertility Penalty and Family-planning Rate	132

List of Figures

1.1	Government Structure in China	27
1.2	Geographical Distribution of Treated Counties in Year 1998	28
1.3	Geographical Distribution of Treated Counties by Year 2004	29
1.4	Geographical Distribution of treated counties by Year 2013	30
1.5	Geographical Distribution of Treated and Control Counties	31
1.6	The Impact of Market Integration: A Case Study of Hangzhou Prefecture	32
1.7	The Impact of Market Integration: A Case Study of Tangshan Prefecture	33
1.8	The Impact of Market Integration: A Case Study of Hengshui Prefecture	34
1.9	The Impact of Market Integration on Economic Development (Approach I) . .	35
1.10	The Impact of Market Integration on Economic Development (Approach II) .	36
1.11	The Overall Impact of Market Integration on Economic Development	37
1.12	The Effect of Market Integration on Reallocation	38
1.13	The Effect of Market integration on Firms' Entry	39
2.1	Provincial Fertility Penalties in Urban China	72
3.1	Four Treatments	106
3.2	Benign Premiums across Treatments	107
3.3	Benign Premiums depending on Malign-game Partner's Action	108
3.4	The Correlation between Action in the Malign Game and How Much shared in DG	109
3.5	Robustness Check - Benign Premium across Treatments	110
4.1	Event Study of the Reform on Geographical Concentration	115
4.2	Robustness: The Impact of Market Integration (Approach II)	116

1.0 Administrative Barriers, Market Integration and Economic Growth: Evidence from China

Along with the reduction in transportation costs in the last two centuries, institutional trade barriers have become increasingly important obstacles to further market integration. This paper examines the impact of a policy reform in China that removed inter-regional administrative trade barriers by incorporating counties into prefectures with a larger market. Using a difference-in-differences approach, I compare incorporated counties, both before and after the reform, to two novel control groups: counties that applied for incorporation but failed and counties that were incorporated several years later. I find that the reform immediately and persistently increased the economic growth of incorporated counties. Several sources of evidence suggest that treated counties experienced relatively rapid growth because they became more integrated into the domestic market. First, using an indirect measure of protection, I find that the reform significantly reduced local protectionism between incorporated counties and their corresponding prefectures. Second, market shares of more productive sectors increased in treated counties following the reform. Third, firms producing tradable goods rapidly entered treated counties. Finally, less profitable firms in treated counties were more likely to exit.

1.1 Introduction

The last two centuries have witnessed an unprecedented increase in trade volumes both between and within nations. Thanks to technological advancements, trade barriers in the form of transportation costs have substantially reduced, contributing to economic growth in both developed and developing countries (e.g. Faber, 2014; Donaldson and Hornbeck, 2016; Storeygard, 2016; Donaldson, 2018). However, institutional trade barriers (e.g., local protectionism) still widely exist, which strongly limit further mar-

ket integration. The election of Donald Trump, Brexit, etc., even mark the rising support for protectionism. Moreover, protectionism not only exists at a national level, but also between localities. Local governments tend to shield local firms from competition, across different regimes. For example, Eyer and Kahn (2017) document that coal states in the U.S. provide large financial incentives to encourage power plants to purchase locally mined coal. While in China, local governments impose a variety of interregional barriers to trade. They provide subsidies to encourage local purchases, make regulations that discriminate against non-local firms, restrict cross-regional trade, and favor local suppliers in procurement to protect firms within their jurisdiction (Young, 2000; Barwick et al., 2017). In this paper, I address the following questions: does the local economy grow faster if local protectionism is eliminated? If yes, what are the underlying mechanisms?

I study the impact of eliminating administrative trade barriers on economic growth by looking at the *incorporating counties into prefectures reform (Chexian Shequ)* in China. Previous studies have extensively documented local protectionism in China (Young, 2000; Bai, Du, Tao and Tong, 2004; Holz, 2009; Long and Wang, 2015; Barwick, Cao and Li, 2017). The incorporating counties into prefectures reform is one of the central government's attempts to eliminate regional administrative trade barriers.¹ The county government, which is at the fourth level of the administrative hierarchy in China, enjoys a high level of fiscal and administrative autonomy. When a county is incorporated into an adjacent prefecture, which is at the third level of the hierarchy, its government loses its autonomy and becomes an agency for the prefecture government. Consequently, the administrative barriers between the county and the prefecture disappear, and the two local markets are expected to integrate.

As pointed out by Donaldson (2015), to empirically test the effect of trade barriers, one needs to solve two challenges. The first challenge is to find a suitable control group that can address the endogeneity issue. Trade barriers may not be the only difference between the treatment and the control group. The second challenge, as proposed by Rubin (1980), is that the control group and the treatment group should not interact; otherwise,

¹ See a news report that the reform in Ningbo Prefecture, Zhejiang Province reduced the regional administrative trade barriers, <http://m.21jingji.com/article/20161011/433d3ea617b2d039480ff9972663cb64.html>.

there may be over-estimation or under-estimation, depending on the direction of the spillover effect.

I adopt a difference-in-differences empirical strategy comparing the economic growth of counties that were incorporated into prefectures to control counties, before and after the reform. Two special features of this reform enable me to overcome the two identification challenges. In the first approach, using information from prefecture governments' five-year city planning books, I construct a list of counties that were chosen by the prefectures to be incorporated, but were not approved by the higher-level governments for various political and geographical reasons. By comparing the treatment group to this *applied-but-failed* group, I adjust for the nonrandom selection of counties by the prefecture. The second approach takes advantage of the fact that among those counties that were successfully incorporated, the timing of the incorporation is arguably exogenous (i.e., no observable characteristics can predict the timing of incorporations). Specifically, I compare counties that experienced the incorporation in year y to counties that would experience incorporation at least τ years later, e.g. in year $y + \tau$ and onwards. The two control groups also partially solve the second identification problem, as the treatment counties and the control counties in both control groups are relatively isolated geographically.

I use GDP and nighttime light intensity to measure economic activity. Nighttime light intensity, gathered from weather satellite recordings, has been increasingly used by economists to measure economic activity in developing countries, especially at the local level (Henderson, Storeygard and Weil, 2012; Hodler and Raschky, 2014; Storeygard, 2016). As nighttime light intensity is not susceptible to political manipulations, it complements GDP as a measure of economic activity.

Using both approaches, I find that the reform significantly increased economic growth in the incorporated counties. Compared to the *applied-but-failed* counties, the reform raised incorporated counties' GDP per capita by 12.6 percent, and nighttime light intensity per square kilometer by 4.4 percent in ten years. Using the second approach, I find that the reform led to a 9.3 percent increase in GDP per capita, and a 5.9 percent increase in nighttime light intensity per square kilometer for counties that experienced

current incorporation, compared to counties that would experience the reform several years later. The effect of the reform is both immediate and persistent. In the first year following the reform, the economic growth of the treated counties already surpassed that of the control counties, suggesting that there were potential trades that were previously prohibited by local protectionism. I also find that the treated counties still grew faster than the control counties ten years after the reform, suggesting that the effects of the reform were persistent.

One potential concern is that the positive effects on treated counties were driven by a migration of economic activity from prefectures to the treated counties. Thus, I also look at the impact of the reform on the overall economic growth. If the reform only induced a transfer, then the overall economic growth of the treated counties and prefectures as a whole should be the same with the control group. However, my results indicate that the reform raised the overall GDP per capita by 6 percent.

Next, using Annual Industrial Surveys, a comprehensive firm panel dataset conducted by the National Bureau of Statistics, I provide four pieces of evidence that are consistent with the underlying mechanism for the reform's positive effect on economic growth being market integration. First, I show that local protectionism - as measured by Bai, Du, Tao and Tong (2004) - decreased. Bai, Du, Tao and Tong (2004) argue that if local protectionism exists, industries with high shares of state-owned enterprises (SOEs) should be less geographically concentrated. The rationale is that local governments tend to protect SOEs to a larger extent than other types of enterprises, as they can capture more resources from SOEs. I first show that the negative correlation between the share of SOEs in an industry and its level of geographical concentration held true at the county level in my dataset. I then show that the reform significantly reduced the negative correlation in the treated counties, compared to the control counties. More specifically, the negative correlation between the share of SOEs in an industry and its level of geographical concentration persisted in the control counties, but not in the treated counties.

Second, I test the impact of the pro-trade reform on the inter-sector reallocations among the incorporated counties. Melitz (2003) show that international trade induces reallocations towards the most productive firms: they tend to enter the export markets

and absorb the shares of the less productive ones. The corresponding hypothesis in my setting is that the reform eliminated the administrative trade barriers, and treated counties should specialize more in industries in which they had comparative advantages. The results show that the production shares of the most productive sector in the treated counties increased by 2 percentage points on average after the reform, which is a 25 percent increase from their original level.² The coefficient estimate of the effect is statistically significant at the 5 percent level.

Third, I show that the reduction in trade barriers enables the incorporated counties to attract more firms producing tradable goods, which suggests that market access in treated counties likely increased. While the impact of the reform on firms producing nontradable goods is uncertain, I observe an immediate increase in the entry of firms producing tradable goods in the treated counties compared to the control counties after the reform. There is no significant increase in firms producing nontradable goods. The immediate entry of firms suggests that the removal of trade barriers is the primary cause of firm entry, since infrastructure improvement would take time to establish.

Fourth, I find that less profitable firms in the treated counties were significantly more likely to exit after the reform. This is consistent with the mechanism that a reduction in trade barriers increases the competition firms face, which forces less profitable firms to exit (Melitz, 2003).

The findings contribute to the literature on domestic trade barriers and their impact on economic growth. Recent studies have primarily looked at the effect of infrastructure or of reducing transportation costs on development (Faber, 2014; Donaldson and Hornbeck, 2016; Storeygard, 2016; Donaldson, 2018). For example, Donaldson (2018) examines the effect of the railroad network in colonial India on agricultural income. He finds that access to railroads raised agricultural income by 16 percent, while access to lines that were planned and surveyed, but never built has approximately zero effect. My work complements the previous work by identifying the effect of local protectionism, another source of trade barriers. My results indicate that eliminating local protectionism as a way of reducing trade costs also contributes to local economic growth. The project also con-

²The sector is defined at the 2-digit industry level.

tributes to this literature by employing two novel control groups to solve the problems of endogeneity and spillover. The latter is possible because my treated counties and control counties are relatively geographically isolated.

My work also adds to the literature on local protectionism, especially in China. Using a province-level regional specialization index, previous studies have provided evidence of local protectionism in China by showing insufficient specialization and an otherwise unexplained correlation between regional specialization and industry characteristics (Young, 2000; Bai, Du, Tao and Tong, 2004; Barwick, Cao and Li, 2017; Holz, 2009). My contributions to this literature are threefold. First, I extend the previous analysis from province level to county level, a much finer unit. The fact that protection practices widely existed at all local levels indicate that local protectionism is deeply rooted in China's political system. Second, even though the existence of local protectionism has been well documented, its impact on economic growth has rarely been discussed. One exception is Barwick, Cao and Li (2017), who show that provincial protection in China's automobile industry led to a substantial consumer welfare loss. I show that eliminating county-level administrative barriers significantly contributed to local economic development. Third, even though it was suggested in previous studies that *the regionally decentralized authoritarian* regime is the cause of local protectionism, I empirically show that a reform re-centralized authority indeed reduced local protectionism.

Lastly, the paper is also related to previous studies on the policy impact of the incorporating counties into prefectures reform (Tang and Hewings, 2017; Liu, Zeng and Zhou, 2019). I contribute to this literature by providing extensive firm-level evidence to show that the mechanism this reform promoted economic growth is that it eliminated institutional trade barriers between the incorporated counties and the prefectures. Moreover, I construct two novel control groups to solve the problem of endogeneity.

The paper proceeds as follows. Section 2 briefly describes the institutional background and two important features of the incorporating counties into prefectures reform. Section 3 describes various data sources used in this study. Section 4 introduces the empirical model, discussing the potential threats to the identification. I present the main results in Section 5. In Section 6, I analyze the channels through which the reform

increased economic development. Section 7 concludes and discusses the policy implication.

1.2 Background

1.2.1 Local Protectionism

The problem of local protectionism has haunted China over the last forty years. Even though the central government has made great progress in transforming its centrally-planned economy to a market-oriented economy, local authorities are still highly involved in market activities across the country. Local governments use various methods to shield local firms from competitors in other regions, which creates inter-regional trade barriers that are harmful to domestic market integration.

It was hypothesized that the *regionally decentralized authoritarian* (RDA) regime is the cause of local protectionism in China (Xu, 2011). Under RDA, the promotion of local officials heavily depends on local GDP growth (Li and Zhou, 2005). Therefore, local governments have strong incentives to protect local firms whose production contributes towards local GDP. In addition, under RDA, local governments have a high degree of fiscal autonomy. Tax revenues collected from firms account for about a third of the local government total revenues.³ Consequently, concerns about revenue also motivate local officials to protect local firms. To further exacerbate local protectionism, RDA also empowers local officials to protect firms in their region. The administrative authority is decentralized under RDA too. Local authorities can make localized regulations, or even laws in some cases, to favor local firms. They also directly manage some state-owned enterprises (SOEs) and can directly invest in some privately owned firms.

As early as in the mid-1980s, it was documented that local governments favored local manufacturers by limiting the mobility of regional low priced-raw materials and reserved them for the local firms (Watson, Findlay and Yintang, 1989; Bernstein and Lü, 2000).

³Data source: Ministry of Finance of the People's Republic of China (<http://www.mof.gov.cn/>).

The central government issued an official order in 1982 to prohibit those practices. After that, two additional decrees were issued by the central government in an attempt to eliminate local protectionism (Holz, 2009). However, protection activities did not disappear; instead, they became more implicit and harder to detect in response.

One means of local protection is a direct subsidy to local, government-owned SOEs. Those SOEs can thus afford to operate even if they face continued losses. The business decisions of those SOEs are also heavily influenced by the local government. The government can make them only buy raw materials and inputs from other local firms. For example, it is widely reported that some local taxi firms only buy automobiles from local car makers (Barwick, Cao and Li, 2017).

Local government can also protect local firms through favorable regulations and selective law enforcement. The government can tailor their regulations so that only the products of local firms can comply with them. For example, the 2018 market entry requirements for new energy automobiles in Shanghai basically exclude all cars, except for those made in Shanghai.⁴ Similarly, in the enforcement of laws or higher-order governments' policies, the local government can choose to enforce them on non-local firms selectively. For example, in a quality control test on electronic bikes conducted by the Liuzhou city government in 2015, all local brands passed the test and all non-local brands failed the test.⁵

As the official orders have not been sufficient to curb local protectionism in the long-run, the central government also experimented with a reform of incorporating counties into prefectures, in the hope that the merger can eliminate the barriers between the two regions. I analyze this reform in this paper.

1.2.2 Institutional Background

It is hard to understand the reform of incorporating counties into prefectures without taking a closer look at the hierarchical structure of the administrative divisions in

⁴<http://news.bitauto.com/hao/wenzhang/629965>

⁵For details, please see the website of Administration for Industry and Commerce, Liuzhou, Guangxi Province (<http://gsj.liuzhou.gov.cn/tzgg/index.html>).

China. At the top of the hierarchy is the central government, followed by provincial-level governments, prefecture-level governments and county-level governments (Figure 1.1). Generally, each level is in charge of overseeing the work carried out by lower levels of the administrative hierarchy.

Within the county level, there are counties and districts.⁶ Even though they are at the same administrative level, the differences between them are substantial. Districts, formally city-governed districts, are subdivisions of a prefecture or a direct-administered municipality. District governments are agencies of the prefecture governments, and are in charge of implementing the policies made by the prefecture governments. On the other hand, county governments are relatively independent of the prefecture governments, and have a high degree of autonomy in both fiscal and administrative aspects.

Counties have more autonomy than districts in both fiscal revenue and spending. After the Tax-Sharing Reform in 1994, the fiscal revenue of county-level governments can be divided broadly into three categories: budgetary revenue, extra-budgetary revenue and off-budget revenue. Budgetary revenue, including value-added tax, business tax, local enterprise income tax and personal income tax, is shared between the county-level governments and prefectural and provincial governments. The proportion of budgetary revenue the county-level governments need to share with prefecture governments is significantly lower for county governments than for district governments. This is partially caused by the fact that the prefecture government can decide its cut of the district governments' revenue, but need to negotiate with the provincial governments to get a higher share of the county governments' revenue. The extra-budgetary revenue includes a variety of non-tax items - paid use of state-owned resources and assets, state-owned capital operating income, fines and confiscations, etc. The county-level governments do not need to share most forms of the extra-budgetary revenue with its upper-level government. However, counties are in charge of collecting more sources of extra-budgetary income than districts, and are still more independent in this category. Also, land con-

⁶Counties are further divided into county-level cities and counties, with the former having slightly more administrative autonomy. However, as the differences between them are small compared to their differences with districts, I do not distinguish between them in most parts of the paper and refer both of them as counties.

veyance fees - revenues from selling the land use rights to a private party, has gained importance in recent years (Han and Kung, 2015). The county governments not only have greater administrative authority to approve land transfers than the district governments, but also share less of the income with the prefecture governments. In terms of fiscal spending, the districts also have less autonomy than counties, as the prefectures control their spending, while the counties can freely use their revenue as long as they meet the regulatory mandates.

As an agency of the prefecture governments, the district governments are also less independent administratively compared to county governments. They have fewer functions than the county governments, as many government functions have been centralized to prefectures. For example, the public security sub-bureau of the districts is managed by the prefecture public security bureau, and thus is not part of the district government. The district governments have no final say in the relevant assessment of the sub-bureau, or the appointment and removal of cadres in that department. However, the county government directly controls the public security bureau in the county. In addition, the district's land resources and planning, industry and commerce, quality supervision, tobacco monopoly, inspection and quarantine are also directly managed by the prefecture bureaus.

As districts have less fiscal and administrative autonomy than counties, transforming counties into districts essentially means authority is re-centralized to the prefectures.

1.2.3 Two Special Features

Two special features of the incorporating counties into prefectures reform enable me to identify the impact of market integration on economic growth.

The first special feature is that, due to the administrative procedure for this reform, there exists a natural control group. For a county to be incorporated into a prefecture, the prefecture first needs to draft a list of counties they plan to merge. Then the prefecture government exchanges information with the county governments to obtain their consent. If the county government and the county's people's congress agree, the prefecture

government will submit a request to the provincial government. If the provincial government approves the request, the proposal goes to the central government, which has the final say on whether a county can be turned into a district and when that can happen.

The control group I employ in the first approach are counties that were chosen by the prefectures to be absorbed, but the application had not been approved by the central government at least until 2013, the last year covered by the study. I collect the list of counties from the prefecture government's five-year city planning books. As illustrated previously, the prefecture governments needed to choose a list of counties they planned to incorporate. The list is usually published in its city planning book. For example, Suzhou, a city in Jiangsu Province, stated that it planned to incorporate six counties in its 1996 planning yearbook: Zhangjiagang, Taicang, Changshu, Kunshan, Wuxian, Wujiang (Suzhou City Planning Manual, 1996-2010). Not all proposed incorporations were approved. Turning a county into a district is against the interest of the provincial government. At about the same time of the incorporating counties into prefectures reform, there is another reform: *sheng zhi guan xian*, which means provinces can directly administer certain counties. A province-directly-administered county shares a larger proportion of its tax revenue with the province and a smaller proportion with the prefecture than a normal county. As argued by Lu and Tsai (2019), there was a strong inter-governmental vertical competition in China's urbanization process. A province has stronger incentives to turn a normal county into a province-directly-administered county or at least keep its county status. Therefore, most incorporation applications were blocked by provincial governments. For example, Suzhou was only able to absorb two (Wuxian and Wujiang) out of the six counties in its incorporation list till 2019. In particular, in exchange for turning Wujiang into a district, it had to allow Jiangsu province to directly administer its richest county, Kunshan, which is also on its incorporation list (Cartier, 2016).

The central government may also reject or significantly delay the incorporation applications for various political and cultural concerns. For example, Shijiazhuang, the capital city of Hebei Province, planned to incorporate Zhengding, Luancheng, Gaocheng and Luquan in both 2001 and 2006. While the other three have been successfully incorporated, Zhengding is still a county as of today. The reason is purely political - Presi-

dent Xi Jinping was once the leader of the Communist Party of China (CPC) in Zhengding county. To preserve his legacy, Zhengding must remain a county even though it has the highest GDP per capita among the four counties in Shijiazhuang's list and its city center is the closest to Shijiazhuang.

Zhengding is not the only example; all the birthplaces of the former general secretaries of the Communist Party of China have not changed their statuses. In addition, counties on the list of National Famous Historical and Cultural Cities are also immune to incorporation to keep their historic county names.⁷

The second feature of this market integration reform is that the timing of the reform varied substantially between and within prefectures. In Suzhou's case, even though both Wuxian and Wujiang were on Suzhou's incorporation list in 1996, the former was incorporated in 2000, while the latter was incorporated in 2012. Changzhou, a city also in Jiangsu Province, planned to absorb Wujin, Liyang and Jintan in its 1996 city planning book. Wujin was successfully incorporated into Changzhou city in 2002, while neither Jintan nor Liyang were incorporated by 2013.

Many factors may affect the timing of the final incorporation. Occasionally, the central government blocked the applications of incorporation for several years. For example, no applications from Jiangsu province were approved between 2005 and 2008. Provincial and central government's transitions may also affect the application process. For example, Wuhan, a city in Hubei Province, started its application for incorporating Huangpi county in 1996, and the application was passed to the central government in 1997. However, due to government transition in 1997, the application was approved in late 1998 and Huangpi was formally incorporated in 1999.⁸

⁷As of 2015, there are 127 National Famous Historical and Cultural Cities in China. For the complete list, please see http://news.youth.cn/jsxw/201509/t20150901_7072311_1.html

⁸Source in Chinese: https://hb.ifeng.com/a/20181227/7127581_0.html

1.3 Data

I use data from many different sources to estimate the effect of the incorporating counties into prefectures reform on economic development. In this section, I present an overview of these data sources and the construction of the variables I use for the analysis.

I construct a panel dataset of counties that experienced the reform and counties that applied for incorporation but hadn't been approved by 2013. The information on treated counties is obtained from the website of the Ministry of Civil Affairs of China.⁹ In order to identify counties that applied for the reform but failed, I collect the list of control counties from the each prefecture's city planning books, which are published every five years. Note that I am only able to identify whether a county applied for the reform or not, but not the year that the county applied for the reform for the first time. Figure 1.5 shows the geographical distribution of incorporated counties and the applied-but-failed counties. As the map shows, almost every province experienced at least one incorporation. Although it's a national reform, one can still observe some geographical patterns. Most of the treated counties were located in eastern and central provinces, while some applied-but-failed counties were scattered in the western provinces. For that reason, I only explore *within-province* variations in treated and applied-but-failed counties throughout the paper.

The nighttime light intensity data is obtained from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) that reports satellite images of the Earth between 20:30 and 22:00 local time. The satellite-year dataset is available for every 30 arc-second pixel (approximately 1 square kilometer). The value of lights are integers values from 0 (no light) to 63. I use the average lights per square kilometer that fall within the boundary of each county/district to proxy for the level of local economic development.

I collect county-level GDP (including the shares of manufacturing and tertiary industry), population, government expenditure and revenue for the years 1995-2013 from

⁹I dropped counties that were merged with a district into a new district because we cannot identify the outcome after the reform. I also dropped counties that experienced the reform but all things remained unchanged for at least five years.

provincial and prefectural statistical yearbooks.

In addition, I observe firms' activities from China's Annual Industrial Survey from 1998-2007. The survey includes information for all state-owned industrial firms and non-state owned firms with prime operation revenue above 5 million RMB. From the annual survey data, I can observe a firm's profit, revenue, employment, industry codes, location at the county level, the year that the firm was founded, the year that the firm exited. Using firm-level data, I estimate firms' productivity, identify the firms' entry and exit, and calculate the geographical concentration of industries at the prefecture-year level for the mechanism analysis in Section 1.6.

For the main empirical estimation, I limit the sample to the time period 1995-2013. The years before 1995 are excluded because a nation-wide tax-sharing reform took place in 1994. The reason I choose to look at years up to 2013 is that the publicly available nighttime lights data is only provided until 2013.

Table 1.1 presents county-level characteristics for incorporated counties and applied-but-failed counties (columns 1 and 2) in the baseline year. There are 71 counties that experienced the reform, and 188 counties that applied for incorporation but failed. Columns 3 and 4 report differences and p-values conditional on province fixed effects. Compared to the applied-but-failed counties, the incorporated counties have more population, food possession per capita, a lower share of rural population, a higher share of manufacturing and tertiary industry, savings and loans, and a student-teacher ratio. Notably, only one of the eleven variables are statistically different at the 10 percent level, and one is at the 1 percent level. Overall, table 1.1 shows that the research design comparing incorporated counties to applied-but-failed counties balances many (although not all) observable covariates, once accounting for average characteristics in the province.

1.4 Empirical Strategy

DID comparing incorporated counties to applied-but-failed counties. In the first approach, I estimate the effects of the reform on a set of county-level outcomes in a

difference-in-differences framework. Specifically, I compare the evolution of outcomes for counties that were successfully incorporated to counties that applied for the incorporation but failed for various political and historical reasons. The equation takes the following form:

$$y_{cpt} = \beta Reform_{ct} + \theta_c + \delta_{pt} + \epsilon_{cpt} \quad (1.1)$$

where y_{cpt} is the outcome variable of interest for county c in province p at year t . Here I look at the log of GDP per capita and log of (1+nighttime lights per square kilometer). The main coefficient of interest is $Reform_{ct}$, an indicator variable that equals 1 for the treated county in years after the incorporation, and 0 for all other cases. θ_c and δ_{pt} are full sets of county and province \times year fixed effects. By conditioning on county fixed effects, the empirical specification absorbs all time-invariant county-specific characteristics. By conditioning on province-year fixed effects, I can control cross-year common changes in provinces that occur even in the absence of the incorporation. Lastly, ϵ_{cpt} is the error term. Standard errors are clustered at the county level to allow for correlation over time within a county.

The difference-in-differences specification relies on the assumption that, in the absence of the reform, the change of outcomes in incorporated counties and applied-but-failed counties before the reform should have parallel trends. I test the validity of this assumption by plotting coefficients of β_τ s and the corresponding 95 percent confidence intervals of the following equation:

$$y_{cpt} = \sum_{i=-5}^{10} \beta_\tau DtoReform_{ct}^\tau + \theta_c + \delta_{pt} + \epsilon_{cpt} \quad (1.2)$$

where $DtoReform_{ct}^\tau$ are indicator variables that equal 1 if year t is τ years after (or before, if negative) the year of incorporation and 0 otherwise; for control counties, it equals 0 in all years. The indicator of “the year before the incorporation” is omitted as the reference year. The coefficients of interest are β_τ s; they represent the differences between treated and control counties in outcome Y , τ years after the incorporation. For GDP per capita and nighttime lights, none of the coefficients before the incorporation are significantly

different from zero. They are also small in magnitude, but they become consistently positive after the incorporation (Figure 1.9).

DID using variation in the timing of incorporation. One natural concern for the first approach is that there might still be some other systematic differences between the incorporated counties and the applied-but-failed counties other than changes in trade barriers. To address this issue, I exploit the variation in the timing of the reforms as an alternative estimation approach. I employ a difference-in-differences strategy that compares economic growth in counties that experience the current incorporation to counties that would experience the reform several years later, before and after the current reform. Even though the treated counties may not be randomly selected, I show evidence that the timing of the incorporation is arguably exogenous (i.e., no observable characteristics can predict the timing of incorporations) in Appendix 4.1¹⁰.

I construct the sample following Deshpande and Li (2019). For each of the 71 incorporations, I take the county that experienced the current incorporation as the treated county, and construct the corresponding control group as counties that would experience incorporation more than five years in the future. The year of the incorporation is set to be year 0. I restrict my sample to event years from -5 to 5 such that the control counties haven't experienced the reform yet. Lastly, I combine all 71 incorporations and build one dataset.

This approach only uses variation in the timing of incorporation, not variation in the event of incorporation (Guryan, 2004; Fadlon and Nielsen, 2015; Deshpande and Li, 2019). The identifying assumption of the difference-in-differences model is that, in the absence of the reform, the change of economic growth in counties that experienced the

¹⁰To examine whether local characteristics predict the timing of the reform conditional on the reform, I limit the sample to counties that haven't been incorporated in that year but will be turned into districts in the future. I estimate the following equation:

$$ReformY_c = \alpha + \Gamma X_c + \epsilon_c$$

where $ReformY_c$ is an indicator variable that equals 1 if the county received the treatment in the year indicated in the column heading. X_c is a vector of county-level characteristics. I include population (lag), manufacturing share of GDP (lag), tertiary share of GDP (lag), ratio of government expenditure to government revenue (lag), ratio of government revenue to GDP (lag), ratio of government expenditure to GDP (lag), log of lights per square kilometer (lag), dummy of provincial capital, dummy of direct-administered municipalities of China.

earlier incorporation would have trends parallel to those experienced incorporation several years later. This empirical approach requires the timing of the incorporations to be as good as random. To this end, I demonstrate that the timing of the reform cannot be predicted by observable characteristics in Table 4.1. This suggests that the timing of the reform is arguably exogenous. To validate this approach, I re-run the test on the parallel-trends assumption with this new control group. Figure 1.10 shows that counties incorporated earlier and counties incorporated later do exhibit parallel trends in the years before the incorporation, both in nighttime lights and in GDP per capita.

To estimate the effects of the reform in regression form using only time variation, I estimate the following equation on the sample:

$$y_{cpit} = \theta_c + \delta_{pt} + \beta_0 Treated_{ci} + \sum_{\tau} D_{it}^{\tau} + \sum_{\tau} \beta_{\tau} (Treated_{ci} \times D_{it}^{\tau}) + \epsilon_{cpit} \quad (1.3)$$

where y_{cpit} is the outcome for county c in province p for incorporation i at year t . The θ_c are county fixed effects, and δ_{pt} are year \times province fixed effects. The variable $Treated_{ci}$ is an indicator equal to 1 if county c is a treated county for incorporation i .¹¹ The D_{it}^{τ} are indicators equal to 1 if year t is τ years after (or before, if negative) the year of the incorporation and 0 otherwise. I cluster standard errors at the county level. The coefficients of interest are the β_{τ} s, capturing the differences in economic growth between treated and control counties τ years after (or before, if negative) the incorporation.

For table estimates, I estimate the following model:

$$y_{cpit} = \theta_c + \delta_{pt} + \beta_0 Treated_{ci} + \delta_0 Post_{it} + \beta (Treated_{ci} \times Post_{it}) + \epsilon_{cpit} \quad (1.4)$$

where $Post_{it}$ is an indicator equal to 1 if year t is after the incorporation.

There's a trade-off in the choice of year gap between the treatment and control group experiencing the reform. While a small year gap is preferable since the control counties are more closely comparable to the treatment counties, it also imposes an upper bound on the time horizon of the analysis (i.e., I can only identify the effect up to that year gap). I present the main results using a five-year gap such that I can identify effects up to five

¹¹Since the same county can appear as a control and a treated county in the data, $Treated_{ci}$ is not co-linear with the county fixed effects

years after the reform. In robustness checks, I demonstrate that the results are robust if I change the year gap to three-year, four-year, six-year, and seven-year gaps (Appendix Figure 4.2).

1.5 Estimates of the Effect of the Reform on Economic Growth

1.5.1 DID Comparing Incorporated Counties to Applied-but-failed Counties

I begin the regression analysis by estimating the difference-in-differences model in equation 1.1. In Table 1.2, columns 1 and 3 are the baseline estimates using a parsimonious specification that includes only county and province \times year fixed effects. I add time-varying county-level controls in columns 2 and 4, including the manufacturing share of GDP, the tertiary industry share of GDP, and the ratio of government expenditure to government revenue.

Figure 1.9 shows the effect of the reform on the log of GDP per capita and lights per square kilometer respectively, based on estimates from equation 1.2. The treated counties did not significantly differ from the applied-but-failed counties prior to the treatment. GDP per capita in the treated counties after the reform increased by 11 percent over the control counties (Table 1.2). The growth rate of GDP per capita shows a gradual increase over 10 years (Figure 1.9, (a)). Consistent with the results on GDP, the average magnitude of the effect on lights per square kilometer is 4.4 percent increase. It takes two years after the reform for lights to reach a stable 10 percent increase.

Overall effect. Next, I explore the overall impact of the reform on the economic growth of the treated counties and the corresponding prefectures as a whole. The unit of observation for overall effect is prefecture-year. Using the same empirical strategy, I compare prefectures that experienced the reform to prefectures that applied for the reform but failed, before and after the reform. For prefectures that have experienced several incorporations, I only focus on the first reform. I estimate a prefecture-level version of equation 1.1.

Figure 1.11 depicts the effect of the reform for the treated counties and the corresponding prefectures as a whole. The treated and control prefectures exhibit parallel trends in GDP per capita and nighttime lights before event year 0, based on estimates from a prefecture-level version of equation 1.2. Table 1.4 shows that the treated counties and prefectures as a whole gain in GDP per capita by 6 percent as a result of the reform. And the estimates are significant at 10 percent level. Notice that the results on nighttime lights are consistent with the results on GDP qualitatively but they are statistically insignificant. The possible reason is that the intensity of nighttime lights is capped at 63 and there is little brightness potential for the already-lighted prefectures.

1.5.2 DID Using Variation in the Timing of Incorporation

To deal with the concern that there might be other systematic differences between incorporated counties and applied-but-failed counties, I exploit only variation in the timing of the reform. Appendix table 4.1 shows that no observable characteristics can persistently predict the timing of incorporations. In the main results, I compare counties that experience the current reform to counties that experience the reform five years later. I demonstrate that the main results are robust to different year gaps in Section 1.5.3. Figure 1.10 shows that the treated and control counties exhibit parallel trends in GDP per capita and nighttime lights prior to event year 0, based on estimates from equation 1.3. GDP per capita and lights per square kilometer increase by 10 percent and 6 percent respectively as a result of the reform in the treated counties (Table 1.3). The estimates for the difference between the two groups of counties are economically large and statistically significant at less than 5 percent.

1.5.3 Robustness

The identifying assumption of the first method is that treatment and control counties would experience parallel trends in outcomes in the absence of the reform. As seen in Figures 1.9 and 1.10, control and treatment counties exhibit parallel trends in the average nighttime lights and GDP per capita prior to the reform. However, it is still possible

that the reform itself is selected by the central government, and this selection could lead to changes in economic development in those counties. The central government might tend to choose richer counties for the reform. To address this concern, I re-run the main results without observations from the direct-administered municipalities of China¹². I find consistent results (in Appendix Figure 4.2) with even bigger magnitudes, suggesting that the positive impact of the reform is not driven by selection.

I use the five-year gap in the second approach as my main results. As a robustness check, I estimate the effects of the reform using alternative year gaps. Appendix Figure 4.2 shows that the results are robust if I change the year gap to a three-year, four-year, six-year or seven-year gap. Specifically, using alternative year gaps, the treated and control counties exhibit parallel trends in GDP per capita and nighttime lights prior to event year 0; and the reform has a significantly positive impact on GDP per capita and lights per square kilometer.

1.6 Mechanism

Results in the previous section indicate that the reform had a significantly positive impact on economic growth, both immediately and persistently. The next question is: what are the mechanisms underlying this positive impact? My hypothesis is the positive impact of the reform is due to the removal of administrative trade barriers between the prefecture and the incorporated county after they have merged into a new prefecture. While I cannot directly observe the reduction in trade costs, I provide four pieces of evidence that are consistent with the hypothesis of a reduction in administrative trade barriers using firm-level data from the Annual Industrial Surveys.

An indirect measure of local protectionism. In the first piece of evidence, I show that local protectionism, as measured by Bai, Du, Tao and Tong (2004), decreased. Bai, Du, Tao and Tong (2004) argue that local governments tend to protect SOEs to a larger extent compared to other types of enterprises since they can derive much more bene-

¹²The Direct-administered municipalities of China are Beijing, Tianjin, Shanghai and Chongqing.

fits from SOEs. So if local protectionism exists, industries with high shares of SOEs are less geographically concentrated. I first show that the negative correlation between the share of SOEs in an industry and its level of geographical concentration still existed at the county level. I then show that the reform significantly reduced the negative correlation in the treated counties compared to the control counties.

I calculate the concentration index developed by Ellison and Glaeser (1997) to measure the geographic concentration of a specific industry within a prefecture. The Ellison-Glaeser index takes the following form:

$$\gamma_{ij} \equiv \frac{G_{ij} - (1 - \sum_c x_{cj}^2) H_{ij}}{(1 - \sum_c x_{cj}^2)(1 - H_{ij})} \quad (1.5)$$

where γ_{ij} is the Ellison-Glaeser index calculated at industry i in prefecture j . $G_{ij} \equiv \sum_c (s_{cij} - x_{cj})^2$ is the raw concentration, where c is any county in prefecture j , s_{cij} is the share of employment for county c in industry i and x_{cj} is the share of total employment of all industries in county c . $H_{ij} \equiv \sum_{ij} z_{ij}^2$ is the Herfindahl index of industry i in prefecture j , with z_{ij} representing the employment share of a particular firm in industry i in prefecture j . The greater the Ellison-Glaeser index, the higher the geographic concentration. It equals zero if all firms randomly pick their locations.

To examine the above hypothesis, I employ a triple-differences framework as follows:

$$\begin{aligned} \gamma_{ijpt} = & \alpha_i + \theta_j + \delta_{pt} + \beta_1 ssoe_{it} + \beta_2 ssoe_{it} \times treatCity_j + \beta_3 (ssoe_{it} \times Reform_{jt}) \\ & + \beta_4 Reform_{jt} + \sum_t \beta_t (ssoe_{it} \times \delta_t) + \epsilon_{ijpt} \end{aligned} \quad (1.6)$$

where γ_{ijpt} is the Ellison-Glaeser index in industry i , prefecture j of province p in year t . $ssoe_{it}$ is the share of SOEs in industry i in year t .¹³ The coefficients of interest are β_1 and β_3 . α_i , θ_j and δ_{pt} are full sets of industry, prefecture and province-year fixed effects.

Column 1 of Table 1.5 firstly establishes the negative correlation found in Bai, Du, Tao and Tong (2004), namely industries with a high share of SOEs were less geographically concentrated. While Bai, Du, Tao and Tong (2004) study this correlation at the provincial level, my results show that local protectionism also exists at the county level. Column 2

¹³I define the share of SOEs as the ratio of total number of SOEs to total number of all firms.

further shows that the reform significantly decreased the negative correlation between SOE share and concentration for treated counties after the reform compared to the control counties. More specifically, in the control counties industries with high shares of SOEs were still less geographically concentrated. And the estimate is statistically significant at 1 percent level. While the negative correlation disappeared in the treated counties after the reform and the magnitude of the correlation is not statistically different from zero.

Inter-sector reallocation. Second, following Melitz (2003)'s analysis on the effect of exposure to trade on inter-firm reallocations, I test the impact of the pro-trade reform on the inter-sector reallocations among the treated counties. The hypothesis is that the reform eliminated the administrative trade barriers, and the treated counties should specialize more in industries in which they have comparative advantages. First, I estimate firms' productivity (or total factor productivity, TFP) using Akerberg, Caves and Frazer (2006)'s method, which is commonly used in the literature. Then I find the most productive sectors at the 2-digit industry level in the baseline year by aggregating TFP at the county-sector-year level. I compare the evolution of production shares of the most productive sectors for treated counties to counties that applied but failed using the following equation:

$$ProductionShare_{sct} = \beta Reform_{sct} + \theta_c + \delta_{pt} + \epsilon_{sct} \quad (1.7)$$

where $ProductionShare_{sct}$ is the production share of sector s in county c of province p at year t . Each sector's production share is defined as the sector's output as a percentage of that county's total output.

Table 1.6 shows that compared to the applied-but-failed counties, the reform caused 1 percentage point increase (a 12 percent increase) in production share for each of the three most productive sectors (p -value=0.034) and a 2 percentage point increase (a 25 percent increase) for the most productive sector (p -value=0.069) in the treated counties. Estimates are statistically significant and economically large. Figure 1.12 further shows that the increase in production share is not driven by pre-trends.

Firms' entry. As the incorporating counties into prefectures reform granted the incorporated counties access to the prefecture market, the market access of firms in the incorporated counties should increase following the reform. One way to test this hypothesis is to look at entries of firms in those counties. As market access was enlarged by the reform, those counties should attract more firms producing tradable goods, but not more firms producing nontradable goods. I construct the firm entry panel for the treatment and control counties each year using data from the Annual Industrial Survey.

Consistent with my hypothesis, Figure 1.13 (based on estimating equation 1.2) displays an immediately entry of firms producing tradable goods in the treated counties compared to control counties. There is no significant increase in firms producing nontradable goods.¹⁴ The different effects on firms producing tradable goods and nontradable goods suggest an increase in market access for the treated counties after the reform. The immediate entry of firms strongly indicates that the removal of trade barriers is the primary cause of firm entries, rather than infrastructure improvement, which takes time to establish. Notice that firms producing tradable goods started to enter even one year before the formal announcement of the reform (Figure 1.13). The possible explanation is that firms (and the treated counties) already knew about the incorporation ahead of the formal announcement by the central government. To further show that the immediate entry of firms is not driven by the positive impact in the year before the reform, I dummy out one year before the reform in Table 1.7 columns 2 and 4. The results are consistent.

Firms' exit. Lastly, I test the hypothesis that a reduction in trade barriers increases the competition firms face, which forces less profitable firms to exit. As the Annual Industrial Survey is conducted annually, I am able to know the exact year a firm exits. More specifically, if the firm is in the dataset in year y , but not in year $y + 1$ and onwards, I then define its exit year as year $y + 1$. Due to the limitation of the Annual Industrial Survey, I can only accurately observe SOEs' exit.¹⁵ From now on I only focus on SOEs' exit. I adopt

¹⁴This result should be interpreted with caution because I do not have many observations of firms producing nontradable goods in my dataset. The annual survey of industrial firms only includes State-owned enterprises and large-scale firms. I cannot rule out the possibility that the reason why I do not find a significant result on firms producing nontradable goods is that they are mainly small-scale private firms.

¹⁵The Annual Industrial Survey include all SOEs. But for non-state firms, only those whose sales are above 5 million RMB are included. There are two potential reasons for me to observe some private firms

a triple-differences approach based on equation 1.8

$$\begin{aligned}
Exit_{icpt} = & \theta_c + \delta_{pt} + \beta_1 profitMargin_{ic,t-1} + \beta_2 profitMargin_{ic,t-1} \times \\
& treatCounty_c + \beta_3 (profitMargin_{ic,t-1} \times Reform_{ct}) + \beta_4 Reform_{ct} \\
& + \sum_t \beta_t (profitMargin_{ic,t-1} \times \delta_t) + \epsilon_{icpt}
\end{aligned} \tag{1.8}$$

where $Exit_{icpt}$ is an indicator equals to 1 if firm i in county c of province p exits at year t . $profitMargin_{ic,t-1}$ is the profit margin, which is defined as profit as a percentage of revenue, for firm i in year $t-1$. $treatCounty_c$ is an indicator equals to 1 if county c is incorporated.

I find that the firms with lower profit margins in the treated counties had a significantly higher probability of closing down after the reform compare to similar firms in the control counties (Table 1.8). Column 1 shows the less profitable firms were in general more likely to exit than more profitable firms, which is consistent with basic economic intuition. Column 2 presents that the probability of exiting for a firm with a low profit margin in the treated counties is three times larger than that of a firm in the control group, likely due to the reform.

1.7 Conclusion

This paper studies the impact of eliminating administrative trade barriers on economic growth by looking at a policy reform in China. Even though we are in the age of globalization, trade is not always free, even within nations. There are multiple administrative trade barriers made by governments in different regimes that impede market integration and, consequently, economic growth. In particular, despite China's massive success in the transition from a centrally planned economy to a market-oriented economy, local governments still widely utilize their administrative autonomy to protect local

disappeared from the annual survey: one is that the private firm exits the market, the other is that the size of the private firm shrinks below the 5-million-RMB threshold. I cannot distinguish between the two.

firms from outside competition. The reform I study in this work is an attempt to solve this problem by incorporating counties into prefectures and thus eliminating barriers between them.

I find that counties that were incorporated into adjacent prefectures experienced higher economic growth after the reform, both immediately and in the long-run compared to the control group counties. Two special features of this reform enable me to solve two challenges in identification for the literature on the impact of trade barriers, namely endogeneity and spillover. The first special feature is that the applications of some counties that were selected by the prefectures to be incorporated were eventually denied by the higher level governments due to multiple political and geographical reasons. Second, among those counties that were successfully incorporated, the timing of the incorporations varied substantially. A simple test suggests that those variations are random. The two control groups constructed using the two features allow me to solve the endogeneity problem and limit the spillover effects to a large extent.

I also provide suggestive evidence that it is the reform's impact on market integration that drives its positive effect on economic growth. I first utilize an indirect measure of local protectionism to test whether the reform indeed reduced protection practices between the prefectures and the incorporated counties. The results indicate that the reform essentially eliminated protectionism in the treated regions, while protection practices still existed in the control regions. Second, I find that the treated counties reallocated their production towards more productive sections after the reform, which is consistent with Melitz (2003)'s analysis on exposure to international trade. Third, I find that more firms producing tradable goods immediately entered the incorporated counties than the control counties, and there were no such trends for firms producing nontradable goods. Fourth, less profitable firms in the incorporated counties became more likely to exit after the reform.

Broadly, this paper also sheds light on the optimal degree of decentralization. The last three decades have witnessed an unprecedented increase in decentralization reforms in both developed and developing countries. Those reforms, and in particular, their effects on economic growth, have drawn much attention from the economics community.

Scholars argued that, theoretically, decentralization can boost economic growth as governments are more efficient at providing public services at a local level, and competition between local governments constrains the Leviathan government. However, decentralization can also lead to a race to the bottom and local protectionism, resulting in less inter-jurisdictional trade and slower economic growth. This paper provides empirical evidence of the negative impact of decentralization. The reform in my work is, in some sense, a centralization reform. It can eliminate trade barriers because authority is centralized into the higher-level government, and thus, the higher-level government can make decisions that are mutually beneficial to both the lower-level government and itself. The positive impact of this reform indicates that even though decentralization can be beneficial to economic growth (Davoodi and Zou, 1998; Zhang and Zou, 1998; Xie, Zou and Davoodi, 1999), it is not always the case and full decentralization may not be optimal (Jin, Qian and Weingast, 2005; Qiao, Martinez-Vazquez and Xu, 2008; Gemmell, Kneller and Sanz, 2013).

1.8 Figures and Tables

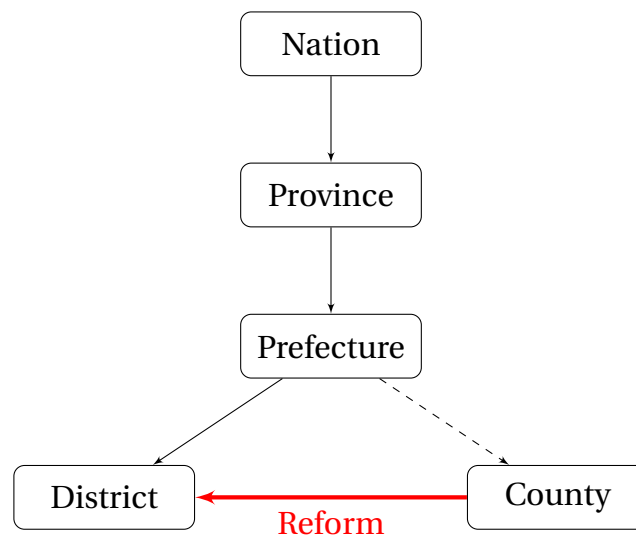
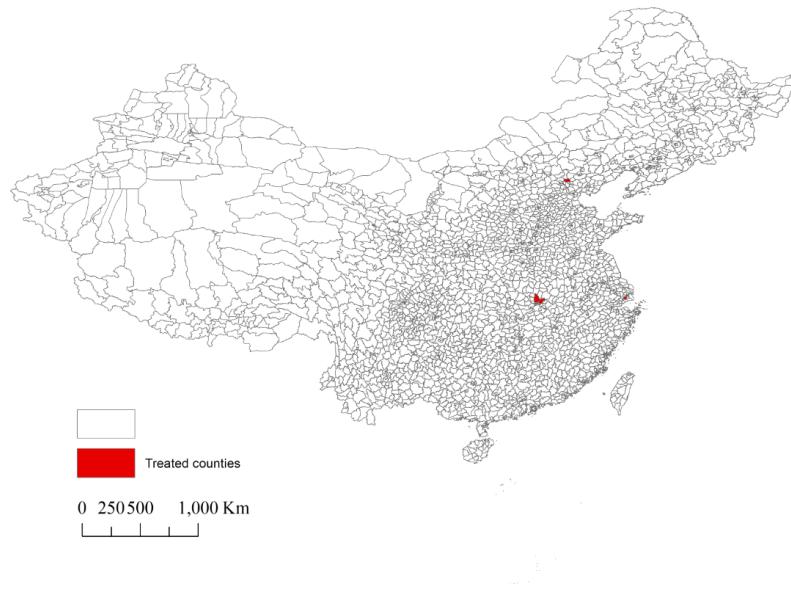
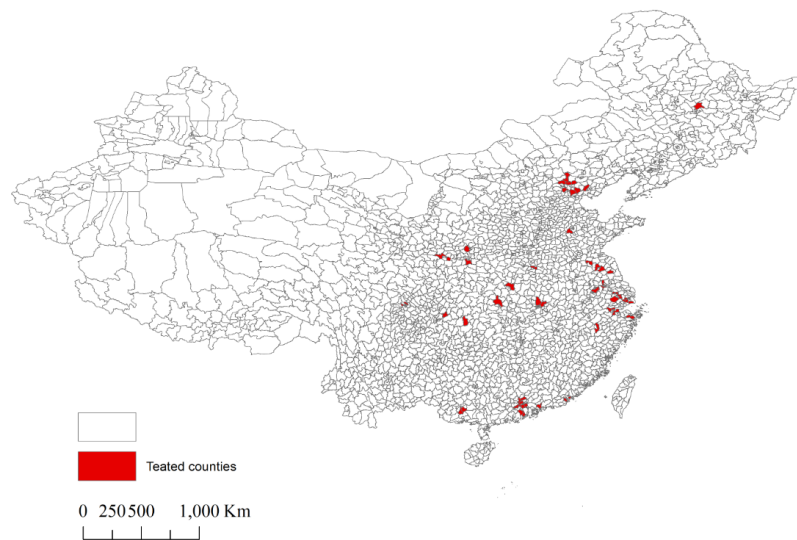


Figure 1.1: Government Structure in China



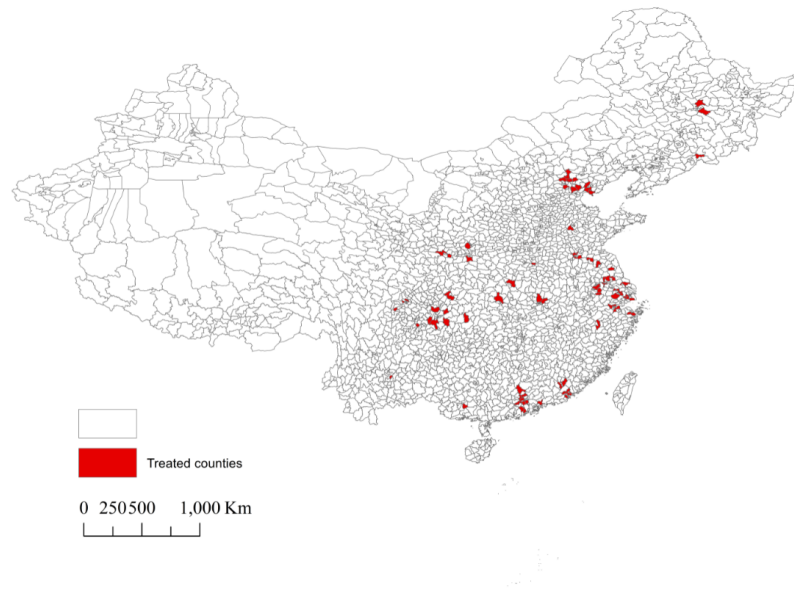
Notes: The map gives the locations of the incorporated counties in year 1998.
Source: Author's mapping based on data from the Ministry of Civil Affairs of China

Figure 1.2: Geographical Distribution of Treated Counties in Year 1998



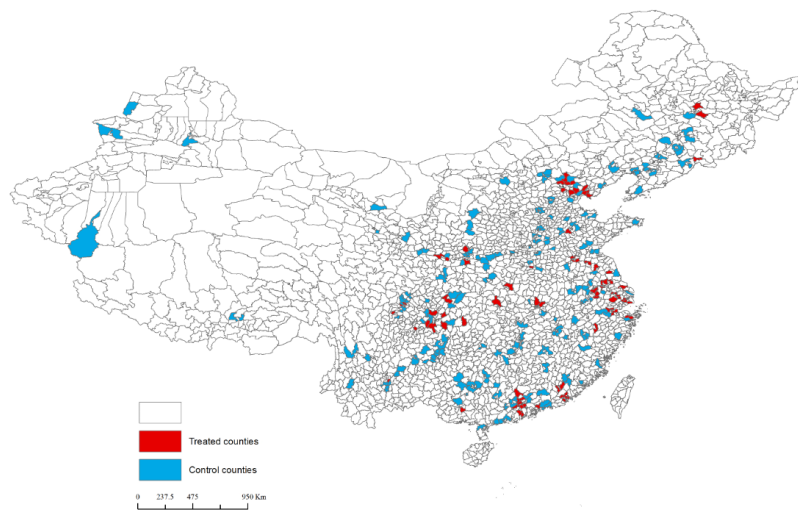
Notes: The map gives the locations of the incorporated counties by Year 2004.
Source: Author's mapping based on data from the Ministry of Civil Affairs of China

Figure 1.3: Geographical Distribution of Treated Counties by Year 2004



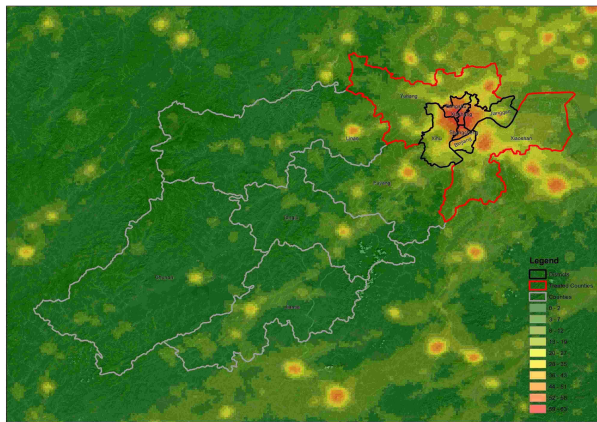
Notes: The map gives the locations of the incorporated counties by year 2013.
Source: Author's mapping based on data from the Ministry of Civil Affairs of China

Figure 1.4: Geographical Distribution of treated counties by Year 2013

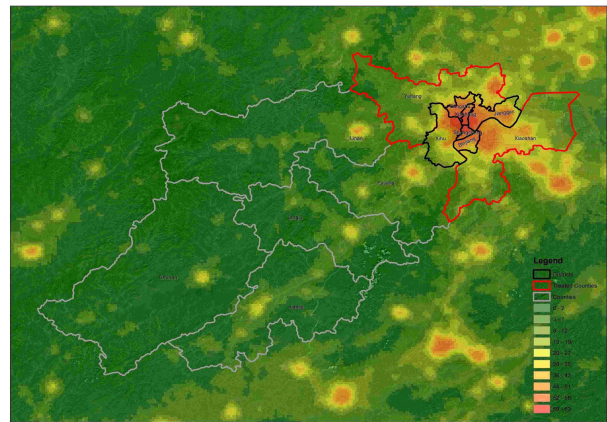


Notes: The map gives the locations of the incorporated counties and applied-but-failed by year 2013.
Source: Author's mapping based on data from the Ministry of Civil Affairs of China and prefectures' city-planning books.

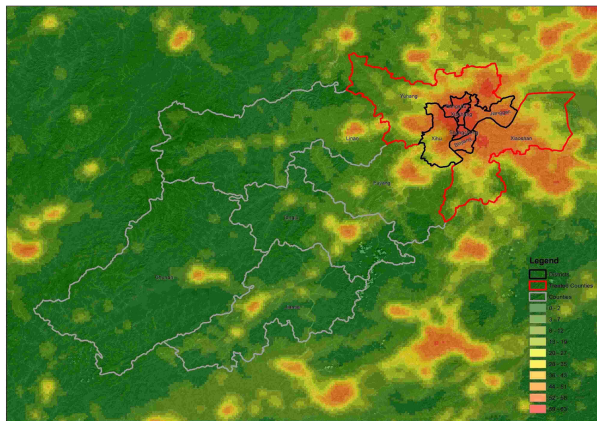
Figure 1.5: Geographical Distribution of Treated and Control Counties



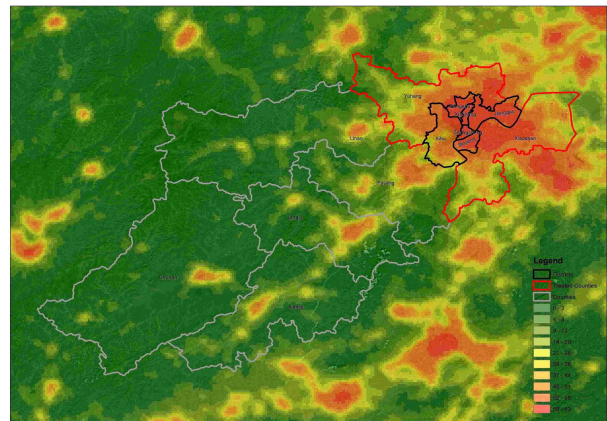
(a) Year 1995



(b) Year 2001



(c) Year 2007

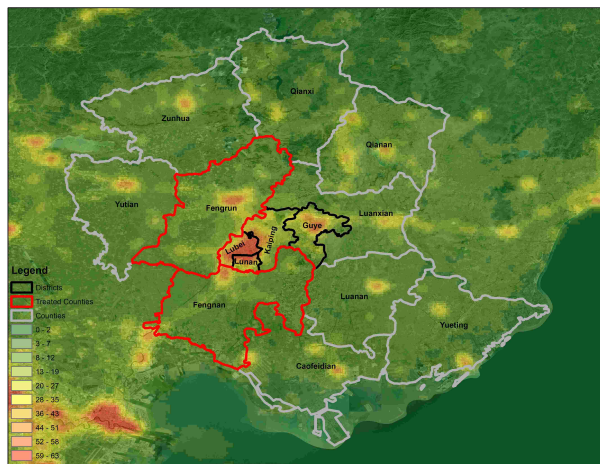


(d) Year 2013

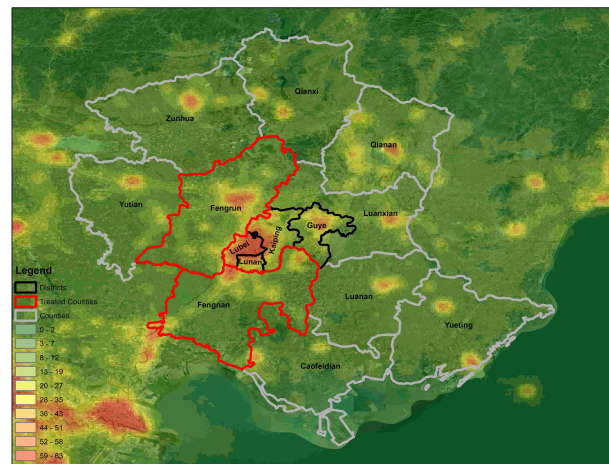
Notes: The map shows the evolution of nighttime lights in Hangzhou Prefecture in Zhejiang Province from 1995 to 2013. Districts are regions with black boundaries. Incorporated counties are regions with red boundaries. Other counties under Hangzhou's supervision are in gray boundaries.

Source: Author's mapping based on data from the Ministry of Civil Affairs of China and the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS).

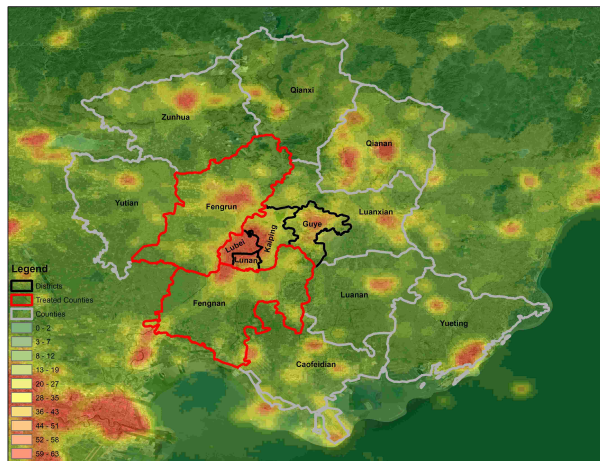
Figure 1.6: The Impact of Market Integration: A Case Study of Hangzhou Prefecture



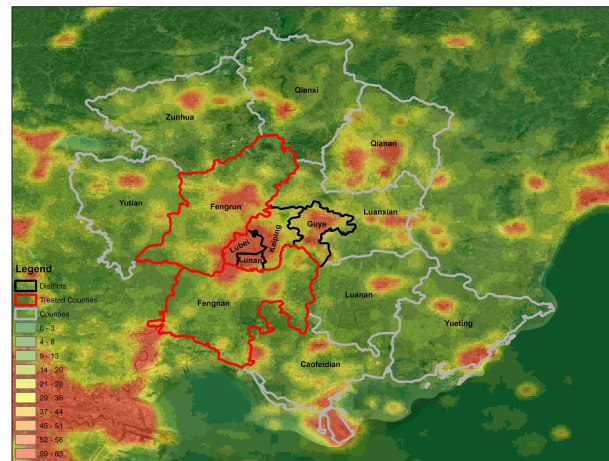
(a) Year 1995



(b) Year 2002



(c) Year 2007

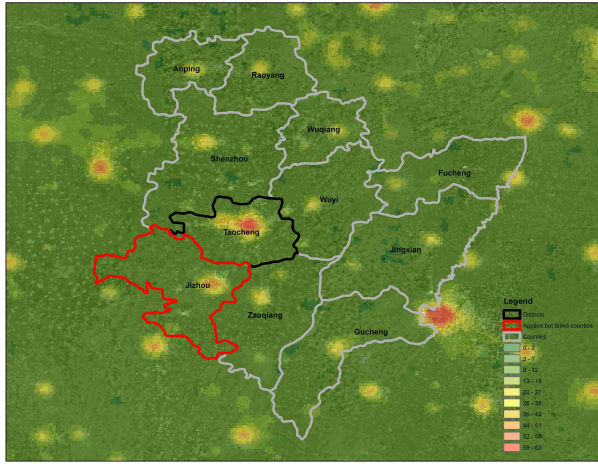


(d) Year 2013

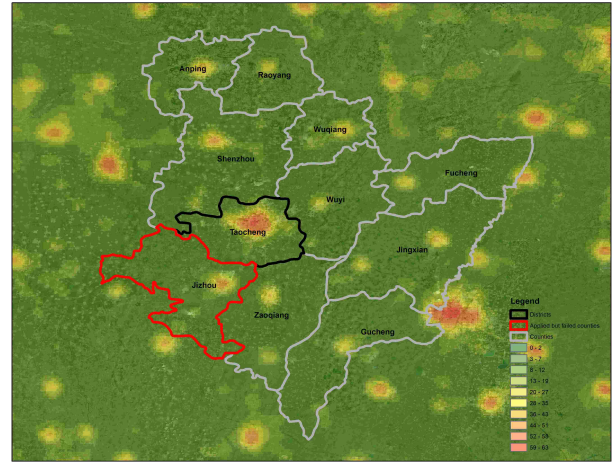
Notes: The map shows the evolution of nighttime lights in Tangshan Prefecture in Hebei Province from 1995 to 2013. Districts are regions with black boundaries. Incorporated counties are regions with red boundaries. Other counties under Tangshan's supervision are in gray boundaries.

Source: Author's mapping based on data from the Ministry of Civil Affairs of China and the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS).

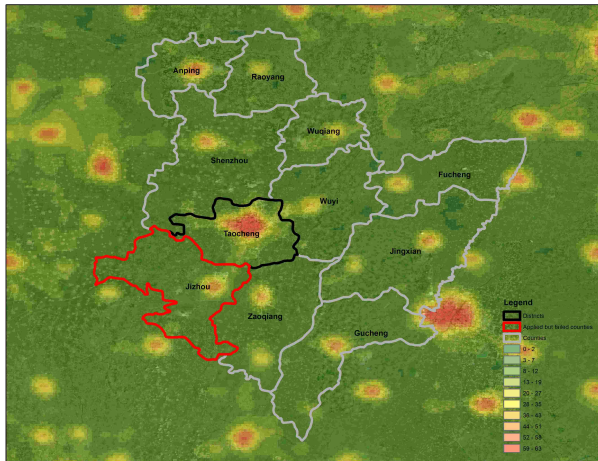
Figure 1.7: The Impact of Market Integration: A Case Study of Tangshan Prefecture



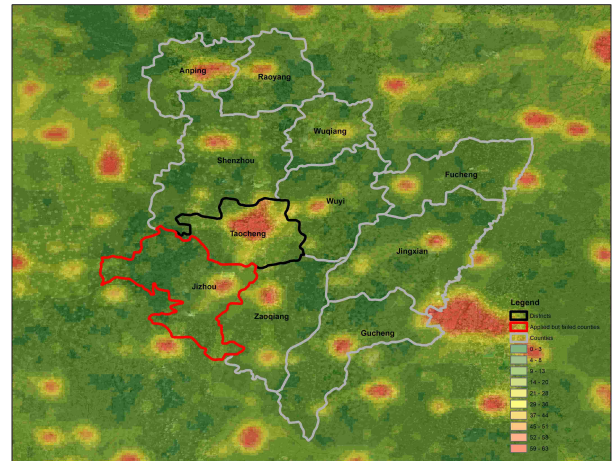
(a) Year 1995



(b) Year 2002



(c) Year 2007

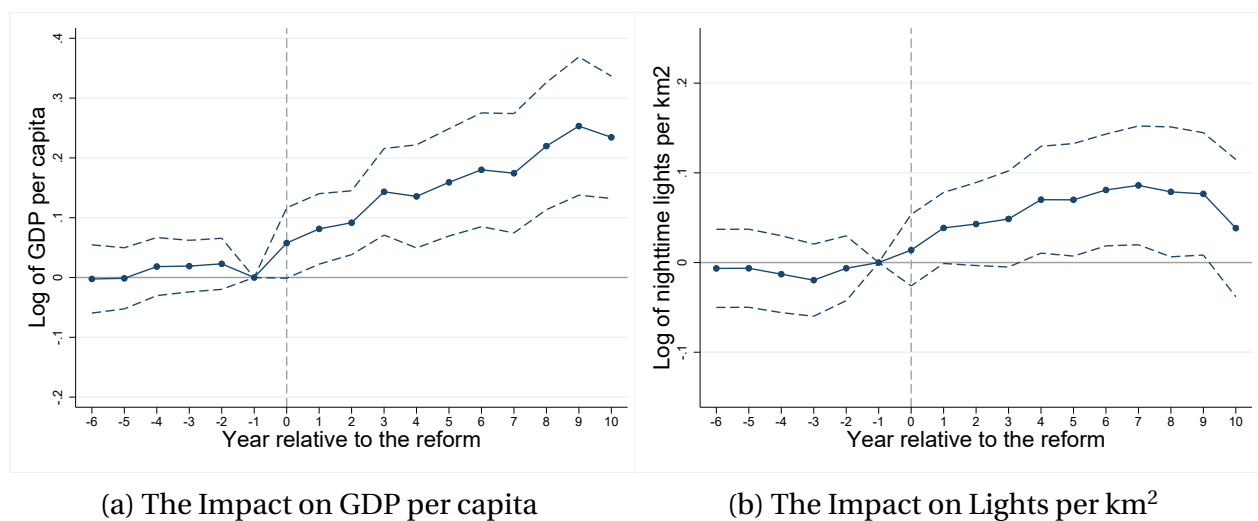


(d) Year 2013

Notes: The map shows the evolution of nighttime lights in Hengshui Prefecture in Hebei Province from 1995 to 2013. Districts are regions with black boundaries. Applied-but-failed counties are regions with red boundaries. Other counties under Hengshui's supervision are in gray boundaries.

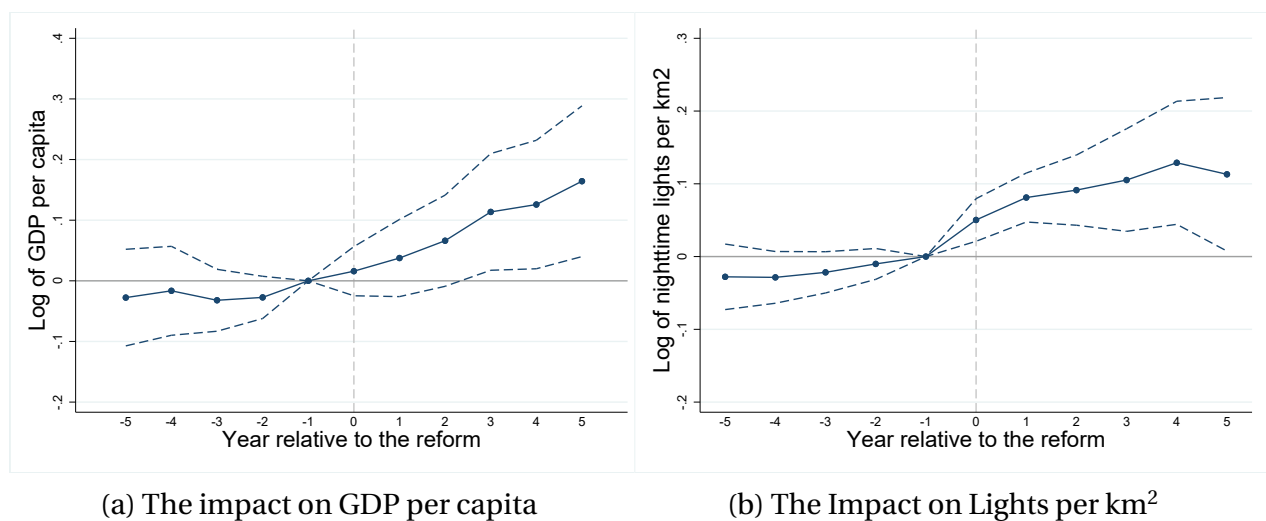
Source: Author's mapping based on data from the Ministry of Civil Affairs of China and the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS).

Figure 1.8: The Impact of Market Integration: A Case Study of Hengshui Prefecture



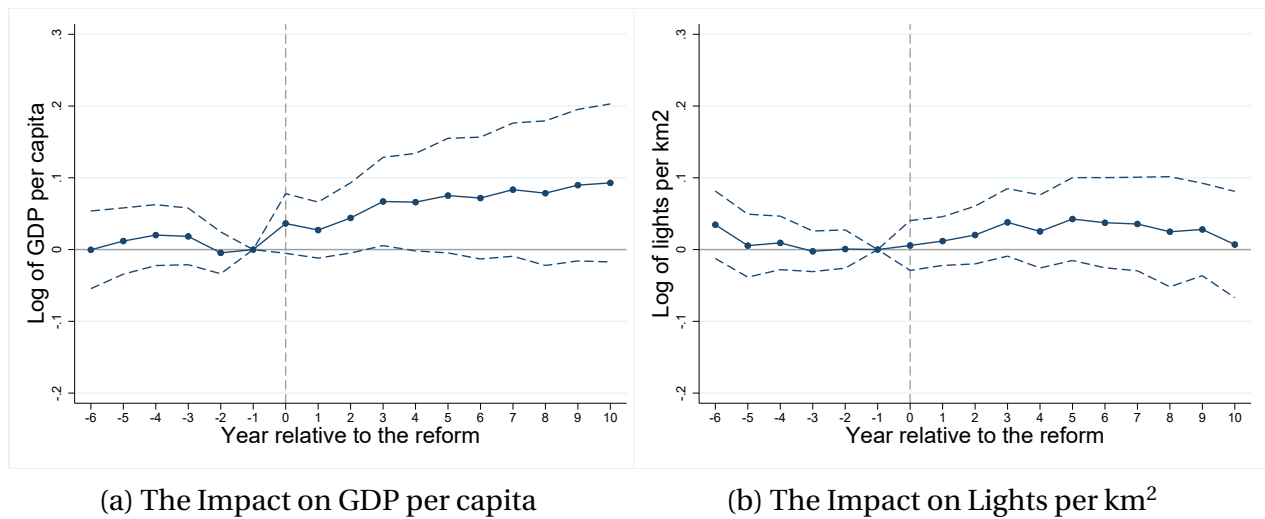
Notes: Figure plots estimates of the effect of the incorporation reform on GDP and nighttime lights in treated counties in the years before and after the reform, based on estimates of coefficients from equation 1.2. The dependent variables are the log of GDP per capita or the log of nighttime lights per km². Dashed line is 95 percent confidence interval for outcomes (solid series).

Figure 1.9: The Impact of Market Integration on Economic Development (Approach I)



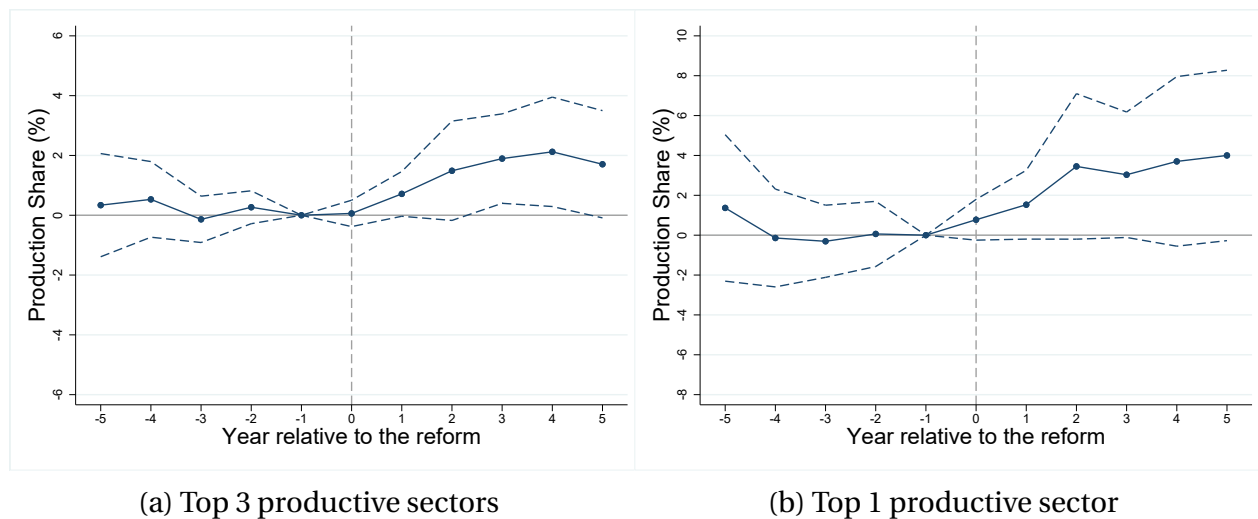
Notes: Figure plots estimates of the effect of the incorporation reform on GDP and nighttime lights in treated counties in the years before and after the reform, based on estimates of coefficients from equation 1.3. The dependent variables are the log of GDP per capita or the log of nighttime lights per km². Dashed line is 95 percent confidence interval for outcomes (solid series).

Figure 1.10: The Impact of Market Integration on Economic Development (Approach II)



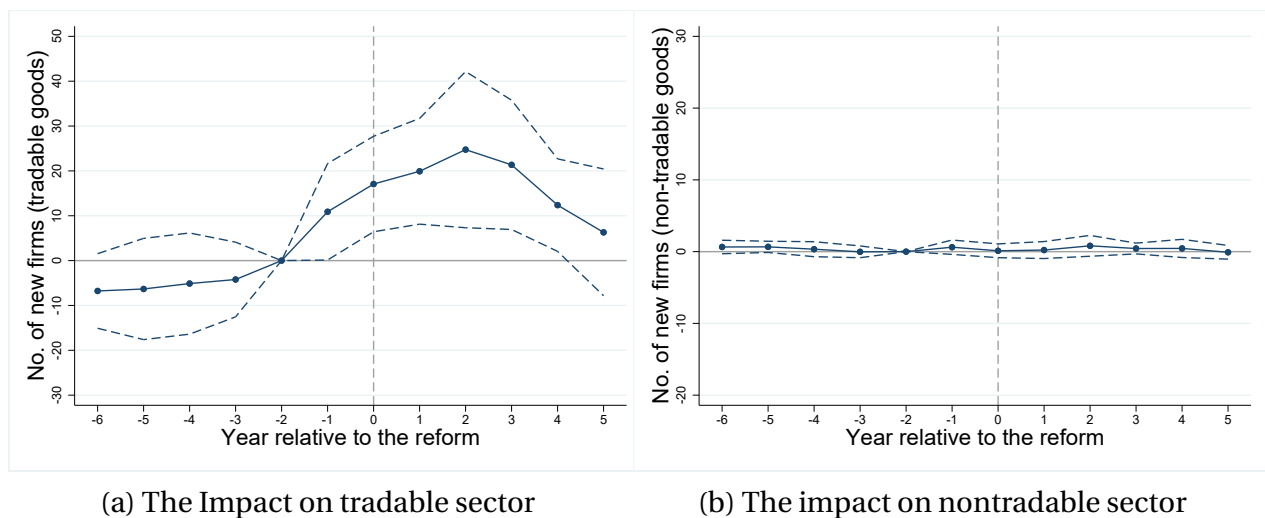
Notes: Figure plots estimates of the effect of the incorporation reform on GDP and nighttime lights in treated prefecture in the years before and after the reform, based on estimates of coefficients from equation 1.2. The dependent variables are the log of GDP per capita or the log of nighttime lights per km². Dashed line is 95 percent confidence interval for outcomes (solid series).

Figure 1.11: The Overall Impact of Market Integration on Economic Development



Notes: Figure plots estimates of the effect of the incorporation reform on production share of the most productive sectors in treated counties in the years before and after the reform, based on estimates of coefficients from equation 1.2. The dependent variable is the production share of the most productive sectors at the 2-digit industry level, which is defined as the output of the sector as a percentage of the county's total output. The sample contains the top three productive sectors in panel (a) and the most productive sector in panel (b). Dashed line is 95 percent confidence interval for outcomes (solid series).

Figure 1.12: The Effect of Market Integration on Reallocation



Notes: Figure plots estimates of the effect of the incorporation reform on the number of new firms in treated counties in the years before and after the reform, based on estimates of coefficients from equation 1.2. The sample in panel (a) are firms producing tradable goods, including all manufacturing firms. The sample in panel (b) are firms producing nontradable goods, consisting of health, education, retail and construction. Dashed line is 95 percent confidence interval for outcomes (solid series).

Figure 1.13: The Effect of Market integration on Firms' Entry

Table 1.1: Summary Statistics (Baseline Year)

	Incorporated counties (1)	Applied-but-failed counties (2)	p-Value (3)
Population (log)	4.170 (.538)	3.809 (.636)	0.053
Share of rural population	.848 (.076)	.868 (.108)	0.131
Share of rural labor participation	.444 (.072)	.432 (.086)	0.801
Food possession per capita	494.8 (220.8)	488.6 (257.8)	0.375
Manufacturing share of GDP	.454 (.090)	.378 (.210)	0.001
Tertiary industry share of GDP	.290 (.077)	.271 (.067)	0.939
Ratio of gov. expenditure to gov. revenue	1.706 (.756)	1.879 (1.607)	0.166
Saving share of GDP	.491 (.238)	.514 (.752)	0.837
Loan share of GDP	.613 (.333)	.615 (.596)	0.598
Students per 10000 people	1564.4 (322.1)	1654.3 (318.3)	0.430
Student-teacher ratio	20.92 (18.59)	19.02 (4.758)	0.792
Number of Counties	71	188	-

Note: This table reports the summary statistics of the treatment and applied-but-failed counties. Column 3 and 4 report differences and p-values conditional on province fixed effects.

Table 1.2: Estimated Effects of the Reform on Economic Growth

Dependent variable	Log of GDP per capita		Log of lights per km ²	
	(1)	(2)	(3)	(4)
Reform	0.122*** (0.035)	0.111*** (0.032)	0.063*** (0.023)	0.041* (0.023)
County-level controls		Y		Y
County FE	Y	Y	Y	Y
Province×Year FE	Y	Y	Y	Y
Observations	4,648	4,458	4,921	4,452
R-squared	0.964	0.970	0.984	0.985
Mean DV	8.968	8.968	1.727	1.730
Std.Dev. DV	0.913	0.913	0.877	0.840

Note: *** p<0.01, ** p<0.05, * p<0.1. The columns presents estimates of β_1 from equation 1.1. All regressions include a full set of county and province×year fixed effect. Robust standard errors are in parentheses, clustered at the county level. The county-level controls include manufacturing share of GDP, tertiary industry share of GDP, ratio of government expenditure to government revenue. Log of population is also included as the county-level control for the results on nighttime lights.

Table 1.3: Estimated Effects of the Reform on Economic Growth: Use
Time Variation

Dependent variable	Log of GDP per capita		Log of lights per km ²	
	(1)	(2)	(3)	(4)
Treatment×Post	0.115*** (0.042)	0.120*** (0.044)	0.095*** (0.027)	0.051** (0.025)
County-level controls		Y		Y
County FE	Y	Y	Y	Y
Province×Year FE	Y	Y	Y	Y
Observations	10,298	10,033	10,344	10,033
R-squared	0.976	0.983	0.988	0.990
Mean DV	8.757	8.784	1.738	1.760
Std.Dev. DV	0.639	0.621	0.707	0.685

Note: *** p<0.01, ** p<0.05, * p<0.1. The columns presents estimates of β_1 from equation 1.4. All regressions include a full set of county and province×year fixed effect. Robust standard errors are in parentheses, clustered at the incorporation level. The county-level controls include manufacturing share of GDP, tertiary industry share of GDP, ratio of government expenditure to government revenue. Log of population is also included as the county-level control for the results on night-time lights.

Table 1.4: Overall Effect of the Reform on Prefectures's Economic Growth

Dependent variable	Log of GDP per capita	Log of lights per km ²
	(1)	(2)
Treatment×Post	0.059* (0.033)	0.012 (0.023)
Prefecture FE	Y	Y
Year FE	Y	Y
Observations	2,795	2,810
R-squared	0.966	0.985
Mean DV	9.385	2.041
Std.Dev. DV	0.913	0.826

Note: *** p<0.01, ** p<0.05, * p<0.1. The columns presents estimates of β_1 from equation 1.1 at prefecture level. All regressions include a full set of prefecture and province×year fixed effect. Robust standard errors are in parentheses, clustered at the prefecture level. Log of population is included as the prefecture-level control for the results on nighttime lights.

Table 1.5: Mechanism: Geographical
Concentration

Dependent variable	Concentration Index	
	(1)	(2)
Share of SOEs	-0.090** (0.038)	-0.125*** (0.048)
Share of SOEs×Treat		0.035 (0.051)
Share of SOEs×Treat×Post		0.116** (0.051)
Industry FE	Y	Y
Prefecture FE	Y	Y
Province×Year FE	Y	Y
Observations	132,444	132,444
R-squared	0.087	0.088
Mean DV	0.215	0.215
Std.Dev. DV	0.594	0.594

Note: *** p<0.01, ** p<0.05, * p<0.1. The columns presents estimates of β_1 , β_2 and β_3 from equation 1.6. All regressions include a full set of county and province×year fixed effects (not reported). In parentheses are standard errors clustered by incorporation. Number of clusters: 152. The industry level controls include number of firms (log), average profit (log) and average employment (log). Number of prefectures: 152 (45 prefectures in treatment and 107 in control). Number of manufacturing industries defined by the four-digit classifications: 424.

Table 1.6: Mechanism: Inter-sector Reallocation

Dependent variable	Production Shares for Most Productive Sectors	
	Top three sectors	Top sector
	(1)	(2)
Reform	1.031** (0.493)	2.102* (1.179)
County FE	Y	Y
Province * Year FE	Y	Y
Observations	5,828	1,871
R-squared	0.343	0.869
Mean DV	8.564	8.550
Std.Dev. DV	13.187	12.826

Note: *** p<0.01, ** p<0.05, * p<0.1. The dependent variable is the production share of the most productive sectors at the 2-digit industry level, which is defined as the output of the sector as a percentage of the county's total output. The columns presents estimates of β_1 from equation 1.1. All regressions include a full set of county and province×year fixed effect. The sample contains the top three productive sectors in column (1) and the most productive sector in column (2). Robust standard errors are in parentheses, clustered at the county level.

Table 1.7: Mechanism: Firms' Entry

Dependent variable	Number of New Firms			
	Tradable sector		nontradable sector	
	(1)	(2)	(3)	(4)
Reform	16.769*** (5.550)	20.161*** (6.346)	-0.028 (0.343)	0.058 (0.402)
One Year relative to the reform		15.740** (6.420)		0.372 (0.459)
County FE	Y	Y	Y	Y
Province×Year FE	Y	Y	Y	Y
Observations	3,063	3,063	594	594
R-squared	0.912	0.913	0.558	0.559
Mean DV	35.092	35.092	1.700	1.700
Std.Dev. DV	63.553	63.553	1.237	1.237

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The dependent variable is the the number of new firms. The columns presents estimates of β_1 from equation 1.1. All regressions include a full set of county and province×year fixed effect. The sample in columns 1 and 2 are firms producing tradable goods, including all manufacturing firms. The sample in columns 3 and 4 are firms producing nontradable goods, consisting of firms providing services in health, education, retail and construction. Robust standard errors are in parentheses, clustered at the county level.

Table 1.8: Mechanism: Firms' Exit

Dependent variable	Dummy for Exit	
	(1)	(2)
Profit margin(lag)	-0.026*** (0.004)	-0.011** (0.004)
Profit margin(lag) × Treat		-0.010 (0.007)
Profit margin(lag) × Treat × Post		-0.027** (0.013)
Industry FE	Y	Y
County FE	Y	Y
Province × Year FE	Y	Y
Observations	247,617	247,617
R-squared	0.060	0.061
Mean DV	0.119	0.119
Std.Dev. DV	0.324	0.324

Note: Profit margin is defines as profit as a percentage of revenue. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The columns presents estimates of β_1 , β_2 and β_3 from equation 1.8. All regressions include a full set of county, industry (at the 4-digit level) and province × year fixed effects (not reported). In parentheses are standard errors clustered by county.

2.0 Responsibility-Shifting through Delegation: Evidence from China's One-Child Policy

There is a growing body of experimental evidence indicating that delegation can foster the shifting of responsibility for unpopular actions from a principal to an agent (Bartling and Fischbacher, 2012). Using the well-known episode of the one-child policy in China (OCP), we provide field evidence for responsibility shifting through delegation. We compare the impact of the OCP on parents who experienced it during 1979-1990 (Phase I), when local governments were the primary enforcer, versus 1991-2015 (Phase II) when the policy enforcement was delegated to the civilians by incentivizing them to report their neighbors' violations of the policy, and appointing cluster leaders to monitor neighbors. Our identification strategy exploits the exogeneity of the gender of the first-born child and argues that parents whose firstborn was a girl were more likely to violate the OCP because of the traditional Chinese "at least one son" preference. Consistent with the predictions of the responsibility-shifting theory, we find that parents who were more exposed to the OCP in Phase II currently trust their neighbors less. The OCP exposure does not undermine trust in local governments. However, parents exposed to the OCP in Phase I currently trust their local governments less. The OCP exposure does not have a significant impact on trust in neighbors in Phase I.

2.1 Introduction

In traditional principal-agent models, principals are assumed to hire agents because the agent either owns private information or has a lower opportunity cost. However, a growing body of experimental literature indicates that the principal-agent relationship might serve to shift responsibility for unpopular actions from the principal to the agent (Bartling and Fischbacher, 2012). In the business sphere, companies use "corporate downsizing consultants" or "firing consultants" to lay off workers on their behalf so

that they do not have to take responsibility. In the political sphere, beginning with Machiavelli, several scholars have proposed that leaders should delegate the enactment of unpopular measures to agents, thus shifting the responsibilities to them (Fiorina, 1986; Vaubel, 1986).

Even though responsibility-shifting through delegation has been found to be quite effective in the lab, field evidence is still missing. In this work, we fill this gap using evidence from China's one-child policy (OCP), which was a large-scale birth control campaign carried out from 1979 to 2015.

The OCP provides us with an ideal setting to study the effect of delegation in shifting responsibility. The OCP is a highly undesirable policy. In 1979, the median family in China had 4.5 children. However, under the OCP, most urban couples were only allowed to have a single child. Violation of the policy led to enormous monetary penalties and non-monetary consequences. Starting in 1991, the enforcement of the OCP was delegated to civilians through mass mobilization. The 1990 census revealed that the government was far behind the target of limiting the population to 1.2 billion in 2000. In response, the government began to involve civilians in the enforcement of the policy. The local authorities created monetary and non-monetary incentives for people to report their neighbors' violations of the policy. Grassroots enforcement organizations were established and civilians were appointed to enforce the policy in their neighborhoods.

The effects of delegation cannot be fully understood unless we simultaneously study what occurs when delegation is not used as a tool for shifting responsibility, which is enabled by the first phase of the OCP enforcement. From 1979 to 1990, the OCP was directly enforced by government officials, especially in urban areas. The family planning commission was established in 1981, and its officials were in charge of enforcing the policy.

Our central hypotheses are: first, when the OCP enforcement was delegated to the civilians in the second phase (1991-2015), parents who were more exposed to the OCP confronted more intense conflicts with their neighbors and trusted them less as a result of responsibility attribution. However, the impact on their trust in local governments is less clear. The local governments could decide the strength of the enforcement of the OCP and its officials were heavily involved in the enforcement process. Thus, it

can be assumed that citizens may have had reasons to hold the governments responsible. The enforcement of the OCP would not affect parents' trust in local governments if responsibility-shifting strategies were effective enough. Second, when local governments were the primary enforcers during the first phase (1979-1990), parents' trust in their neighbors should not have been affected as the neighbors were not involved in the enforcement. However, parents who were more exposed to the OCP are thought to have trusted local governments less as they were the sole party to be held responsible.

The strength of the OCP enforcement varied across provinces and time, depending on local economic conditions, the demographic setting, and other political concerns. We construct an individual-level measure of the OCP exposure in urban areas from 1979 to 2015. In the second phase, we used the average fertility penalty rate a person faced within five years after the arrival of his or her first child as the measure. The fertility penalty was the amount of monetary punishment one needed to pay for an above-quota birth. The fertility penalty not only proxied the strictness of the OCP enforcement as suggested by the previous literature (Ebenstein, 2010; Huang, Lei and Sun, 2015), but it also positively correlated with the financial incentives the local governments could provide for the informers. For example, an informer in Chongqing in 2009 could receive 5 percent of the fertility penalty paid by the victims. We count the penalty rates in the years after the birth of the first child because citizens were restricted by the OCP only after having their first child. The five-year interval is selected because, according to the Chinese Census, most people's second child arrived within five years after the first child.

The fertility penalty data is largely unavailable for Phase I. Therefore, while keeping the construction of the OCP exposure measure the same, we replace the fertility penalty with the rate of family planning, which was defined as the percentage of couples who were of fertility age and had taken birth control measures, like vasoligation, sterilization, and intrauterine devices (IUDs). In the 1980s, voluntary birth control was quite rare. Thus, the family planning rates measures how successful the local governments were in enforcing the OCP. In fact, the family planning rate was one of the main indicators used by higher-order governments to evaluate the performance of the local birth-planning commissions.

Responsibility is the variable of interest in this study, and trust is our proxy variable for responsibility. A lower level of trust is a sign of being held responsible. In particular, we focus on citizens' trust in their neighbors and local governments, which are measured by the 2016 China Family Panel Studies (CFPS-2016).

The main empirical strategy of the study exploits the exogenous variation in the gender of the first child.¹ Given the deep-rooted belief that one family needs *at least one son* to maintain the family's lineage, parents whose firstborn was a girl were more likely to violate the OCP by trying to have a second child, unlike parents whose firstborn was a boy. Thus, parents with a firstborn daughter were more exposed to the OCP.²

Our results are consistent with the predictions of the responsibility-shifting theory. In Phase I when enforcement of the OCP was delegated, exposure to it leads to a significantly larger reduction in trust in neighbors for those whose firstborn child was a girl than for those whose firstborn was a boy. However, the OCP exposure in this phase does not undermine trust in local governments. Furthermore, the coefficients on trust in neighbors are significantly different from coefficients on trust in local governments at the 5-percent level. On the contrary, in Phase I without delegations parents with a firstborn girl lose more trust in local governments when facing stronger exposure to the OCP compared to parents with a firstborn boy. At the same time, the OCP exposure does not have a significant impact on trust in neighbors in Phase I. Moreover, the coefficients on trust in local governments are significantly different from the coefficients on trust in neighbors at the 1-percent level.

We contribute to the lab responsibility-shifting literature by providing, to the best of our knowledge, the first field evidence (Coffman, 2011; Hamman, Loewenstein and Weber, 2010; Bartling and Fischbacher, 2012; Oexl and Grossman, 2013). In a typical lab

¹Gender of the first child was previously used by Li and Wu (2011) to measure the bargaining power of women within the household, and by Wei and Zhang (2011a) to vary the strength of competitive saving motive.

²A natural concern, given our identification strategy, is the possibility that the gender of the first child was not perfectly exogenous due to pervasive sex selection practices, such as selective abortion in China. However, multiple sources of evidence suggested that sex selection rarely happened with the first child (Li, Yi and Zhang, 2011; Wei and Zhang, 2011b; Li and Wu, 2011). We also directly test the correlations between the gender of the first child and all the control variables in the empirical analysis (Table 2.3). We do not find a significant correlation.

setting, the principal in the control group chooses between a fair allocation and an unfair allocation that benefits herself at the cost of a recipient. In the treatment group, another option is added: she can delegate the allocation choice to an agent whose interest is aligned with hers. The authors find that when the unfair allocation is chosen by the principal, the recipient who is adversely affected is willing to costly punish the principal harshly. However, if the task is delegated to the agent and the unfair allocation is chosen, then the principal receives a much smaller punishment and the agent is punished. In this paper, we replicate the main results in a large-scale field setting.

By confirming the effectiveness of delegation in responsibility avoidance, this paper also provides a rationale for mass mobilization. By encouraging people to fight against each other, governments can avoid the responsibility for implementing an unpopular policy, a cost they must bear if they need to do all the work themselves. Even though mass mobilization is widely observed in authoritarian regimes, it has been overlooked in economics. The only exception is Lichter, Loeffler and Siegloch (2015). Using county-level data of the number of informers in the 1980s in East Germany, the authors show that higher levels of government surveillance led to lower levels of political trust in post-reunification Germany. The key difference is in their scenario, even though the number of informers was quite large, they were still contracted government employees. The governments were still responsible for their actions, while in our setting, the informers were mobilized civilians.

In a broader context, our findings contribute to the growing literature focusing on conflict and trust. Nunn and Wantchekon (2011) identify a persistent impact of the historical slave trades on current trust levels within Africa. Rohner, Thoenig and Zilibotti (2013) document causal effects of ethnic conflict on trust and ethnic identity using multi-level data from Uganda. Chen and Yang (2019) study the causal effect of the Great Chinese Famine (1958-1961) on the survivors' and the subsequent generation's distrust in the government. We provide a clear mechanism of the effect of conflict on distrust. We show how mass mobilization, a measure popular among governments during a social conflict, affects people's interpersonal and institutional trust.

Lastly, this study adds to the literature studying the impacts of the OCP. The conse-

quences of the OCP range from economic growth (Li and Zhang, 2007) to sex ratio imbalance (Ebenstein, 2010; Li, Yi and Zhang, 2011), female education (Huang, Lei and Sun, 2015), and to competitive saving motive (Wei and Zhang, 2011a). We add to this literature by showing that the enforcement of the policy incurred a hidden cost to the civil society by lowering people's interpersonal trust.

Our paper proceeds as follows: Section 2.2 describes the historical background, institutional setup, and important features of the OCP. Section 2.3 describes the various data sources used in this study. Section 2.4 introduces our identification strategy and empirical model. We present the main results in Section 2.5. We provide a discussion of alternative explanations in Section 2.6. Section 2.7 concludes and discusses the policy implication.

2.2 Background of the One-Child Policy

China's one-child policy is credited with dropping the total fertility rate from 2.81 in 1979 to 1.51 in 2000 (World Bank). The Chinese have long favored large families. Total fertility exceeded six births per mother throughout the 1960s (Banister, 1991). In the 1970s, after two decades of explicit encouragement of population growth, policymakers in China enacted a series of measures to curb population growth. The OCP was introduced in 1979 and began to be formally phased out in 2015. Under the OCP, most urban couples were only allowed to have a single child. However, the regulations varied among regions. Provincial governments localized the state fertility policy due to the diversity of demographic and socioeconomic conditions across China. As the main instrument for enforcing the OCP, financial penalties also varied across provinces.

The government employed various methods to enforce the OCP. Parents after the first or second birth were required to have an intrauterine device (IUD) inserted or undergo sterilization. The sterilization rate, which is defined as the percentage of women of reproductive age who underwent sterilization, increased from 21 to 35 percent between 1979 and 1999 (Scharping, 2013). At times, the government used a more draconian method

- induced abortions of unauthorized pregnancies - as a “remedial measure making up for contraceptive failures” (Scharping, 2013). For above-quota births, the government mainly used fertility penalties to enforce the OCP. Depending on the province of residence and time period, the fertility penalty for an unauthorized child equaled 10 to 25 percent of a family’s annual income for 7 to 14 years (Serrato, Wang and Zhang, 2016). In urban areas, other forms of punishments were also widely used. For example, people employed in urban units were threatened with the denial of health and welfare benefits, bonus payments, lack of job promotions or even demotions.

From the 1990 census, the central government found a large number of “excess” births during 1986-1990. In order to achieve the 1.2 billion population limit for the year 2000, the central government began a stricter enforcement of the OCP. In May 1991, the central government started a “one-vote-down” (yi piao fou jue) campaign. The chief officials at each administrative levels were made personally responsible for achieving the birth-control targets. If targets were not achieved, the chief officials were not promoted while some even lost their jobs. Second, in December 1991, the central government began to take birth-control performance into the regular performance evaluations of government officials.³ Local authorities increased the grants for the program and gave it stronger political support. Enforcement of the program was also strengthened. The campaigns resulted in a huge jump of penalties for one unauthorized child and were succeeded by a substantial decline in fertility.

In response to the “excess” births found in the 1990 census, the central government launched a mass mobilization campaign. This movement was clearly stated by Tieying Li, who was a member of the Central Politburo of the Communist Party at that time. He said “No one is allowed to give birth beyond the birth-quota, and let the masses watch each other,” in an internal speech on April 21, 1990. The rationale for this campaign was twofold. First, the local cadres lacked the necessary information about who was pregnant and whether it was above quota or not. There were not enough local officials to

³The Performance Evaluation System for government officials is an important component of the government personnel management in China. Cadres take the evaluations very seriously. The evaluation result is one of the most influential factors affecting their career appointments, promotions, transfers, and removals (Wang, 2013).

monitor every woman of childbearing age, and those parents who planned to give unauthorized births intentionally hid from them. The government thus relied on people who were close to the pregnant women to provide the information. Second, there were not enough local officials to enforce the policy even if information was provided. Enforcing the OCP was not only about sterilizations and forced abortions. The cadres also launched propaganda campaigns to promote the idea of one child and carried out the so-called “Three Examinations,” which checked women for the use of contraceptive rings, pregnancy, and illness four times (or more) a year. An expanded crew, with a limited budget, was needed within a short time frame, which made asking citizens to be “volunteers” highly attractive.

Two measures were taken to mobilize the masses. First, to deal with the information asymmetry, the citizens were encouraged and incentivized to report unauthorized pregnancies and births of their neighbors, coworkers, and relatives to the authorities. The incentives provided by the government were mainly monetary.⁴ The payments varied across regions and were linked to fertility punishments. For example, in Chongqing in 2009, the informer could be awarded 5 percent of the fertility penalty paid by the victims, which was the equivalent of one-month salary.⁵ In addition, the informer’s identity was kept confidential and there was no record of punishments for a false report.

Second, to deal with the ever-increasing workload, more at-will employees and “volunteers” were recruited by the government. They carried out most of the detailed work instead of government officials. They were generally seen by people as neighbors.

In order to discuss the recruitment and duties of those workers, facts about local management in urban China are needed. The family-planning commission was the institute in charge of the enforcement of the OCP. Its lowest level was located in the county

⁴In some cases, reporting could also lead to career rewards, and failing to report could sometimes cause collective punishments. The career rewards for denunciation were salient when there were competitions between colleagues. Public sector employees who were caught violating the OCP would not be promoted in most cases and might even lose their jobs. When there was a quota for promotions, the hidden career rewards could be huge. Collective punishments were collected when someone was aware of an OCP violation but failed to report it. During the years when the OCP was most fiercely implemented, one worker’s violation of the OCP could lead all of her or his coworkers to lose a significant part of their income in some state-owned enterprises. (<http://wap.sciencenet.cn/blogview.aspx?id=749707>)

⁵<http://www.chinalawedu.com/lvshi/AAA635949214532/58912.shtm>.

government, which worked with residents' committees to achieve the birth-planning targets. The residents' committee (sometimes also translated as the neighborhood committee) is the lowest level of urban administration in China. According to the Chinese constitution, it enjoys a high degree of autonomy and is named the "self-government organizations of the masses." It is allowed to recruit at-will employees and pays them independently. Although the residents' committees is formally the lowest organizational entities, there are entities one level below them – the residents' small groups (jumin xiaozu). They may comprise a neighborhood or just an apartment building. The small groups are a means of an internal organization and do not have a legal status of their own. Their leaders are called "cluster leaders." They are appointed by the residents' committee, but they are not on the government payroll.

Enabled by the revenue collected from the fertility penalty in the 1990s, the residents' committees recruited many at-will full-time and part-time OCP enforcers. For example, Huangjiapu, a residents' committee with a population of 500 in Shanxi Province, had 15 full-time at-will employees tasked with family-planning matters in its peak (Fong, 2016). As those people were paid by the residents' committees but not the higher authority, they were not entitled to the social welfare benefits enjoyed by government employees, and thus were not cadres in people's eyes. They were often seen as neighbors instead. The recruitment of the OCP enforcers still failed to solve the labor shortage. Consequently, the cluster leaders were mobilized to enforce the policy. They were tasked with keeping track of households' reproductive habits and reporting those details to the local family-planning commission. These leaders were also seen as neighbors. In addition to the cluster leaders, state organizations such as the military, public schools, and hospitals had their internal family-planning units, as did state-owned enterprises.

These neighborhood-level staffs and cluster leaders were the basic building blocks of China's OCP machinery. According to a report issued by the national family-planning commission, while there were only half a million full-time employees combined in the central and local commission, there were about 1.2 million neighborhood-level birth-planning staff and more than six million cluster leaders who were mobilized to enforce the OCP (Fong, 2016).

2.3 Data

To estimate the effect of delegation on trust, we use measures of trust from the China Family Panel Studies (CFPS) and individual-level OCP exposure data from Scharping (2013) and other sources. The CFPS is a nationally representative panel survey conducted by the Institute of Social Science Survey at Peking University. We introduce our measurement of different types of trust and exposure to the OCP in section 2.3.1 and section 2.3.2, respectively.

2.3.1 The Measurement of Trust

The primary outcomes of interest are citizens' trust in neighbors and local governments, which are measured by the CFPS-2016 survey. The questions asked in the survey are translated as follows: Please rate to what degree you trust your neighbors? Similar questions were also asked on trust in local government officials, not in local government *per se*. This ensures that when we compare trust in neighbors and trust in local governments, we are comparing two types of trust, whose objects are both people. In that sense, we are not comparing people's trust in two totally different domains. One may also be concerned with the validity of categorical trust measures. However, Johnson and Mislin (2012) provide experimental evidence that trust, as measured by the World Values Survey, is positively correlated with experimentally measured trust. The questions asked in the CFPS are very similar to those in the World Values Survey. From the summary statistics (Table 2.1), we can see that people generally trust their neighbors more than the local governments.

2.3.2 The Individual-Level OCP Exposure

Our key explanatory variable is an individual-level measure of the exposure to the OCP: the average fertility penalty an urban resident faced during the five years after the arrival of his or her first child. More precisely, for an individual i living in urban areas of province p whose first child was born in year t , his or her exposure to the OCP is measured

by the mean value of the penalty rate in province p from year $t + 1$ to year $t + 5$.

We use provincial fertility penalties for one unauthorized child to measure the strictness of the OCP enforcement. As displayed in Figure 2.1, fertility penalties varied across provinces and across time. At the provincial level, as documented by Scharping (2013), there were three forms of fertility fines. The first form was a percentage deduction from the wage over several years. For example, in February 1980, Guangdong province ratified a fine of 10 percent of income from each parent for 14 years for an unsanctioned birth. The second type of fines was levied as a lump sum payment based on annual income. For example, Shanghai reported in 1992 that an unauthorized birth carried an immediate payment of three years of household income. The third form was a certain amount of immediate payment regardless of household income. For example, from 1995 to 2000, Guangxi ratified the fine as an amount between 2,000 RMB and 50,000 RMB. Following Ebenstein (2010), we transform all three types of fines into percentages of household income (see Appendix 4.2.1 for more details). Regarding the time trend, fines increased over time, but the timings of the changes were quite different among provinces. Our measure of the OCP exposure incorporates both the provincial and time variation.

We improve the measure of the OCP enforcement in the literature in two ways. First, we modify the formula for transforming the fines into percentages of household income to reflect the rapid and unbalanced growth in Chinese economy in the last 30 years.⁶ We also extend the penalty data from 2000 to 2015.

Second, we construct an individual-level measure of the OCP exposure. In our measure, only penalty rates implemented after parents had their first child count. The reason is that individuals were restricted by OCP only after having the first child. China Census data show that the interval between the birth of most couples' first and second child was no more than five years (Scharping, 2013). Therefore, the strongest impact of the birth control policy fell on young couples during the five years following the birth of their first child. That is why we only count the rates in that 5 years. However, as shown in the next section, changing the interval to 4 or 6 years does not qualitatively alter our results. This measure exploits the individual variations in the timing of the first child's birth. Many

⁶More details of our calculation are provided in Appendix 4.2.1.

factors play a role in the timing of having a child. The chance that two women of the same age give birth to their first child at the same time is low.⁷

Sample selection. For our empirical estimation, we limit our sample to individuals who completed all the CFPS-2010, CFPS-2012, and the CFPS-2016 surveys. We further limit our sample based on two criteria: (i) individuals that resided in the urban sector, and (ii) individuals who gave birth to their first child between 1979 and 1985 or between 1991 and 2010.

We restrict our sample to urban households because we do not have a valid measure of OCP exposure in rural areas. Compared to urban areas, rural areas are farther away from administrative centers. It was much harder for the higher level government to make sure that the local cadres closely followed the provincial policy. Furthermore, it was also difficult for the local cadres to collect information on the households' annual income, on which the fertility penalty was based and calculated. In practice, the fertility penalty was often set to be the same for all households in the same village for the sake of lacking information. Additionally, some low-income families could not afford the massive amount of penalties. In urban areas, this was not a problem as the penalties could be collected on a monthly basis and be directly deducted from the salary. However, the cash flow of the rural residents was not as stable as that of their urban counterparts. Rural cadres often chose to collect as much as the low-income family could afford on a lump-sum basis. Hence, there was a lot of randomness in the enforcement of the OCP in rural areas.

For criteria (ii), the reason we choose to look at people who gave birth to the first child after 1979 is it is the year the OCP started. We restrict our sample to individuals who gave birth to their first child before 2010 because the OCP was formally phased out in 2015. We exclude people who gave birth to the first child after 2010 from our sample as they only experienced OCP during part of the five-year interval after the birth of their first child. For a similar reason, we exclude people whose first child arrived between 1986 and 1990

⁷One may worry about the possibility that parents rationally choose the timing of the first birth to enjoy looser OCP enforcement. However, that is unlikely for two reasons. First, the provincial OCP enforcement policy was mainly made by the provincial government and was highly unpredictable for people without a special connection to the provincial government. Second, to choose a looser OCP enforcement, one would need to plan the timing of having the first child, the timing of the second, and foresee the enforcement strength for several years. Even if it could be done, it might be too costly to implement.

as we cannot tell whether they experienced Phase I OCP or Phase II OCP.

2.4 Empirical Strategy

2.4.1 Identification Strategy

Our main identification strategy exploits the exogenous variation in the gender of the first child. A deep-rooted belief in the Chinese culture is that each family needs *at least one* son to maintain the lineage. Consequently, urban couples whose first child was a girl were more likely to violate the OCP by trying to have a second child than parents whose first child was a boy. If so, they were more exposed to the OCP penalties. To validate our empirical strategy, we provide evidence that the gender of the first child is truly exogenous. We then show that the propensity to give birth to a second child was higher among couples whose first child was a girl.

One potential concern is that the gender of the first child is not exogenous because sex selection was a widespread practice in China. However, there were few sex selections performed for the first child. Sex selection practices can be performed either prenatally (sex selective abortion) or postnatally (for example, female infanticide). The bulk of sex selections in China took place prenatally as the accessibility of sex-selective technology improved (e.g., Edlund, 1999; Das Gupta, Jiang, Li, Xie, Woojin and Bae, 2003). Chen, Li and Meng (2013) provide evidence of few sex-selective abortions on the firstborn child using data from the Chinese Children Survey. They show that for first pregnancies, the sex ratio at birth (males/females) was close to being natural.⁸ In some years, it was close

⁸According to (Wilson and Hardy, 2002), the natural sex ratio at birth is estimated at 106 boys to 100 girls. A range of natural and environmental factors may affect the natural sex ratio. For example, Mathews et al. (2005) provide extensive evidence to show that the following factors are relevant. The first factor is the age of the mother: mothers aged 25 to 35 had babies with a sex ration of 1.05 on average. But the sex ratio ranged between 0.94 and 1.11 for mothers who were below the age of 15 or above 40. The race was also an important factor: the ratio was 1.05 for the white non-Hispanic population, 1.04 for Mexican Americans, 1.03 for African Americans and Indians. It is the highest (1.07) for mothers of Hawaiian, Filipino, Chinese, Cuban, and Japanese ethnicity. The maximum value went as high as 1.14 over the 62-year study period. However, for the results on racial differences, whether those differences were purely driven by nature or whether social factors also played a role is still an open question.

to the average natural rate of 1.06 and was never higher than 1.14. Moreover, the abortion ratio (the number of abortions/number of children born) was smaller than 0.05 percent for the firstborn child. The positive correlation between sex ratio and abortion rate was driven mostly by second and higher order pregnancies.

Our statement that the gender of the first child is exogenous is confirmed and complemented by the nationwide census data of China in 1990, 2000, and 2005.⁹ As documented in Table 2.2, the ratio of boys and girls for the firstborn child was also close to being natural. More specifically, the male/female sex ratios were 1.052, 1.071, and 1.024 in the three waves of the census.

To further check whether the gender of the first child is truly exogenous, we run a regression of the gender of the first child on all other control variables that are used later in the empirical analysis. Table 2.3 reports the results using both a linear probability model and a probit model. People did not endogenously choose the gender of the first child based on the fertility penalties in previous years. Also, none of the other control variables is statistically significant in explaining the gender of the first child.

The exogenously determined gender of the first child affects Chinese parents' later childbearing behaviors. Here, we define the "at least one son preference" as always preferring to have a new child when there is no boy in the children profile. This specific definition of "son preference" indicates that people whose first child was a girl are more likely to try to have a second child than people with a firstborn boy. Even though trying to have more than one child could lead to OCP penalties, for some, the potential benefits of having a son outweighed the cost for some. However, for people whose first child was a boy, the potential benefits of an additional child vastly diminishes and they may find it too costly to violate the OCP. Therefore, we should expect parents whose first child is a girl to be more exposed to the OCP.

This "at least one son" belief is deeply rooted in Chinese culture. Confucianism, as the "state religion" in ancient China, is still strongly influencing the Chinese people in

⁹Since there is no information on children who moved out of home, we limit the sample to mothers and children who are mostly likely to be living in the same household. More specifically, we restrict our sample to married women aged 21-40 who had their first child after the one-child policy was introduced. We further restrict their matched children aged between 0 and 18.

modern times. In Confucian philosophy, filial piety is one of the four virtues. It means to be good to one's parents, which requires ensuring male heirs. Mencius or Mengzi, who is the most famous Confucian after Confucius himself, once said: "There are three forms of unfiliality, and bearing no heirs is the worst" (Chan, 2002).

Evidence from census data suggests that the "at least one son" preference was quite persistent in modern China. Table 2.2 illustrates the probabilities of having more than one child after having either a firstborn boy or a firstborn girl using census data. It suggests that, for instance, 49 percent of those with a firstborn daughter had more than one child, while for those with a firstborn son the number was 36 percent in the 2000 census. We consider this gap to be large as people who were not affected by the OCP were also included in the census.

2.4.2 Empirical Specification

Let i index individuals, c index birth cohorts, and p index provinces. We model an outcome of interest y_{icp} , which could be trust in neighbors or local governments. Our key independent variable of interest is $1stChildpenalty_{icp}^{1-5}$, which we define as the five-year mean value of the penalty rates in province p after individual i had his/her first child.

$$\begin{aligned}
 y_{icp} = & \sum_c \alpha_c + \sum_p \delta_p + \beta_0 1stChildpenalty_{icp}^{1-5} + \mathbf{X}'_{icp} \gamma + \beta_1 1stChildGirl_{icp} \\
 & \sum_c \alpha_c \times 1stChildGirl_{icp} + \sum_p \delta_p \times 1stChildGirl_{icp} + \mathbf{X}'_{icp} \gamma \times 1stChildGirl_{icp} \quad (2.1) \\
 & + \beta_2 1stChildpenalty_{icp}^{1-5} \times 1stChildGirl_{icp} + \epsilon_{icp}
 \end{aligned}$$

where $1stChildGirl_{icp}$ is the dummy variable for the gender of the first child. $1stChildGirl_{icp}$ is equal to 1 if individual i 's first child was a girl. β_0 captures the effect of OCP exposure on trust for individuals whose first child was a boy. The main coefficient of interest is β_2 , which captures the differential impact of OCP exposure on trust of people with a firstborn girl and trust of people with a firstborn boy. \mathbf{X}'_{icp} is a vector of observable characteristics for individual i of birth cohort c in province p . Here, we include income and education attainment level. α_c and δ_p are full sets of birth cohort and

province of current residence fixed effects. By conditioning on province fixed effects, our empirical specification absorbs all time-invariant province-specific trust characteristics. By conditioning on cohort fixed effects, we can difference out cross-cohort changes in trust that would occur even in the absence of OCP. Lastly, ϵ_{icp} is the error term. We cluster our standard errors at the province level to allow for correlation over time within a province. Due to the smaller number of clusters in this case (25), we implement a wild cluster bootstrap-t procedure (Cameron, Gelbach and Miller, 2008) for improved inference. We also present the corresponding p-values in our tables.

2.5 Results

2.5.1 The Effect of OCP Exposure between 1991 and 2015 on Trust

In Table 2.4, we present the regression estimates from equation 2.1 on the two trust outcomes: trust in neighbors and trust in local governments. In columns 1 and 3, we present the baseline estimates using a parsimonious specification that includes only province and birth cohort fixed effects. We add individual-level controls as our preferred specification from equation 2.1, including education attainment and family income per capita. Columns 2 and 4 indicate that our results are consistent across different specifications. The remaining discussion focuses on our preferred specification.

Regarding trust in neighbors (column 1), the OCP exposure significantly lowers trust in neighbors for couples whose first child was a girl, but not for couples whose first child was a boy. The estimated differences between the two groups (β_2), are significantly different from 0 at the 1 percent level and economically large. As the gender of the first child is arguably exogenous, this result suggests that the effect we find is causal. The parameter estimates indicate that a one standard deviation increase in the OCP exposure is associated with 0.244 standard deviation more reduction in trust in neighbors for people whose first child is a girl than for people whose first child was a boy.

Next, we proceed to examine whether the OCP exposure affects people's political

trust. Due to the political sensitivity of eliciting trust in the central government in mainland China, we are only able to measure citizens' trust in local governments explicitly. Columns 3-4 of Table 2.4 present the corresponding estimation results. One can see that there is no significant impact of OCP exposure on trust in local governments regardless of the gender of the first child. Also, there is no significant difference between the two groups of people. The estimate of coefficient β_2 is small and not significantly different from 0. This finding is particularly striking because people blame their neighbors for turning them in, but they do not blame the governments who initiated the harm. One can also see that our results are consistent if we use wild bootstrap p-value.¹⁰

2.5.2 The Effect of OCP Exposure in the 1980s on Trust

In the previous section, we showed that when the OCP enforcement was delegated, more severe enforcement significantly undermined people's trust in neighbors but not trust in local governments. Nevertheless, another question arises: if local governments enforced the policy by themselves instead of delegating it to the neighbors, would the people blame local governments but not the neighbors? The early ages of the OCP enforcement enables us to study this possibility. The policy came into effect in 1979, but the masses were not mobilized until 1991. Thus, parents whose first child was born between 1979 and 1990 were exposed to the OCP but were not hurt by their neighbors. Looking into their trust in local governments and neighbors can help us answer the above question.

Ideally, we want to keep our analysis consistent and use the fertility penalty to measure OCP severity. However, 11 out of 31 provinces had not established their fertility penalty policies in 1984. Some provinces introduced their first fertility penalties in 1988. We use the family-planning rate instead as the measure of the OCP severity in this period. The rate of family-planning is the percentage of couples who are in fertility age and have taken birth-control measures such as sterilization, IUDs, and birth-control pills. Voluntary birth control was quite rare in the 1980s; the family-planning rates measured

¹⁰We conduct a triple-difference analysis to further show that the effects of OCP exposure on trust in neighbors and trust in local governments are significantly different from each other at the 5 percent level.

how successful the local governments were at enforcing the OCP. It was one of the main indicators used to evaluate the performance of the local family-planning commissions.

One limitation of this approach is that the local governments have strong incentives to over-report the family-planning rate in order to achieve its birth-control targets. However, as long as the degree of over-reporting is the same across province and time (which is likely as local cadres faced similar incentives when reporting the family-planning rate), our measure of the OCP exposure is still valid. Even if this assumption fails to hold, the degree of over-reporting is still orthogonal to the gender of the first child and thus our identification strategy can handle this problem.

Consistent with our individual-level measure of the OCP exposure, we use the five-year average family-planning rate an urban resident experienced during the five years after the arrival of his/her first child. More precisely, for an individual i living in urban areas of province p whose first child was born in year t , his/her exposure to OCP is measured by the five-year average family-planning rate in province p from year $t + 1$ to year $t + 5$. To further validate the family-planning rate measure, we cross-check the correlation between fertility penalties and the family-planning rate when both have data in the same year, the two measures are significantly correlated at the 5 percent level (Table 4.13).

We can see from Table 2.5 that when the policy was solely implemented by government officials during the 1980s, the results are quite the opposite: citizens blame local governments, but not neighbors. For people whose first child was a girl, a one standard deviation increase in family-planning rates leads to a 0.38 standard deviation more decrease in trust in local governments relative to people whose first child was a boy. The estimates for the difference between the two groups of people (β_2) are statistically significant at the 1 percent level. In contrast to the results on people's trust in local governments, OCP exposure has no significant impact on trust in neighbors regardless of the gender of the first child and the difference between the two groups of people is not significant.¹¹

¹¹Using a triple-difference analysis, we show that the effects of OCP exposure on trust in neighbors and trust in local governments are significantly different from each other at the 5 percent level.

2.5.3 Mechanism

One important concern is the question of whether our results are driven by systematic differences between couples whose first child was a girl and couples whose first child was a boy beyond the OCP exposure.

To distinguish between the two explanations, we look at the differences between people who only had one child and people who had more than one child. If the difference between people whose firstborn was a boy and people whose firstborn was a girl is indeed the difference in OCP exposure, then we should expect that for people who had more than one child the differential effect of OCP on trust should disappear. The reason is that urban couples who had more than one child, regardless of the gender of the first child, all violated the OCP. Thus their exposure to the OCP was the same regardless of the gender of the first child. However, if the differential effect of OCP exposure on trust are driven by other systematic differences between the two groups, then we should expect the results not to be affected by how many children they had.

Table 2.6 and Table 2.7 show that our results are driven by people who had only one child. In both Phase I and Phase II, our results remain consistent if we restrict our sample to those who only had one child. However, there's no significant difference in trust in neighbors or local governments between people whose first child was a daughter and people whose first child was a son if they had more than one child. This result is consistent with our hypothesis but not with other explanations of our main results.

2.5.4 Robustness Checks

One potential problem of our work is that the gender of the first child is not perfectly exogenous. We argue that, if anything, such selections are likely to work against the results we find. Here, the couples whose first child is a girl are the treatment group and the couples whose first child is a boy are the control group. Due to the sex selections favoring boys, some couples who are in the treatment group are mistakenly identified as the control group, which makes it harder for us to observe a difference between the two groups. Furthermore, Table 4.4 lists the number of male and female first births across provinces

from June 1999 to June 2000 in urban areas according to the 2000 census. Since the sample sizes are too small to calculate a precise sex ratio, we construct a one-tailed t-statistic to test whether the calculated sex ratio (column 4) is statistically different from the biological sex ratio¹². We find that six provinces' sex ratios at first birth are significantly higher than the normal sex ratio at the 5 percent level: Beijing, Jiangsu, Jiangxi, Hubei, Guangdong, and Guangxi. To further validate our results, we drop observations from the six provinces and replicate the specifications in Table 2.4. Estimation results, reported in Table 4.5, indicate that dropping them does not affect our findings.

Although we have argued that it is unlikely for parents to endogenously choose the timing of the first birth to enjoy looser OCP enforcement, we perform another robustness check using an alternative measure of the OCP exposure: the minimum penalty (family-planning rate) in the five years after people had their first child after 1991 (before 1991). We re-estimate our baseline equation with the alternative measures of OCP exposure. The results are unaffected (Table 4.6 and Table 4.7 in Appendix).

We next examine the robustness of the OCP exposure measure. Recall that we used the five-year average fertility penalties (or family-planning rates) after people had their first child. Our results are quite consistent when we use a four-year (Table 4.8 and Table 4.10) or six-year (Table 4.9 and Table 4.11) average fertility penalties (or family-planning rates).

2.6 Discussions

We showed when the enforcement of the OCP, an extremely unpopular policy, was delegated to civilians, there was a significant decline in people's trust in their neighbors but not in their trust in local governments. However, when government officials enforced the OCP, people's trust in government was undermined. The results are consistent with the responsibility-shifting effect of delegation. In this section, we address whether we can attribute the results to the delegation. In what follows, we explore three alternative

¹²Appendix 4.2.2 further shows the detailed construction of the one-tailed t-statistic.

interpretations of our results: the local governments are held responsible in the second phase but people are unwilling or dare not report their mistrust; people do not blame the local governments but they do blame the central government in the second phase; or the local governments in the second phase performed better than the governments in the first phase and their good performance canceled out the responsibility.

2.6.1 Dare Not To Report Mistrust

An important concern regarding the absence of a significant impact of OCP exposure on trust in local governments in the second phase is that people try to express politically correct views. To address this possibility, we present the distributions of responses to the question of trust in local governments (Table 4.12) in this period. From the broad range of answers to the question, an indication is that respondents are likely not attempting to provide “correct” responses. If there is such a “correct” response, then one would expect that 10 is the correct value. However, in fact, people’s trust in local governments is not abnormally high. The mean value is smaller than the mean value of trust in neighbors in our sample.

Also, we find a significant negative impact of OCP exposure on trust in local governments in Phase I when government officials were in charge of enforcing the policy, which further suggests respondents are willing to report their mistrust in local governments. Thus it is unlikely that people’s fear to report mistrust in local governments drives our Phase II results.

2.6.2 Trust in Central Government

The second concern about our estimates is that people do not hold the local governments responsible in the second phase because people treat the local government as merely the policy enforcer, not the policymaker. So it is important to see whether there is any effect on trust in the central government in Phase II. Due to the political sensitivity of eliciting trust in the central government in mainland China, we are only able to measure trust in local governments. Even though we can potentially use trust in local

governments to cautiously extrapolate trust in the central government (Cantoni, Chen, Yang, Yuchtman and Zhang, 2017), one might wonder how big the correlation is between the two trust measures. As an additional check, we construct two proxies for trust in the central government as the dependent variables.

The first proxy for trust in the central government is an indicator variable where 1 represents the individual is/was a member of the Communist Party. Choosing to join the Communist Party *per se* reflects one's political attitude and belief.¹³ The second proxy is the difference between the number of days accessing political news through television and the number of days accessing political news through the Internet. It is a well-established fact that all television broadcasters in China are the "mouthpiece" of the Party (Zhao, 1998; Shirk, 2011). While Internet censorship is much looser than media censorship, Internet censors focus their attention on silencing speech that may generate collective action, rather than criticism of the government (King, Pan and Roberts, 2013). Moreover, Internet users can browse international news channels that are not provided on television. So if one does not trust the central government, she is more motivated to access the political news from a less censored media source - the Internet. We calculate the difference in the media source to get rid of the variation in how much one cares about political news. The bigger the difference, the higher the level of trust in the central government.

From Table 2.8 across all columns, we do not find any significant difference in the impact of the OCP exposure on trust in the central government between people whose first child was a girl and people whose first child was a boy. This suggests that people not only do not blame the local governments, but they also do not blame the central government either.

¹³This measure certainly has its shortcomings. One issue is that people may choose to become a member of the Communist Party simply due to career concerns. Another concern is that people may lose their party membership due to violations of the OCP, which can also explain a negative correlation between OCP exposure and the probability of being a party member.

2.6.3 Better Performance in the Second Phase

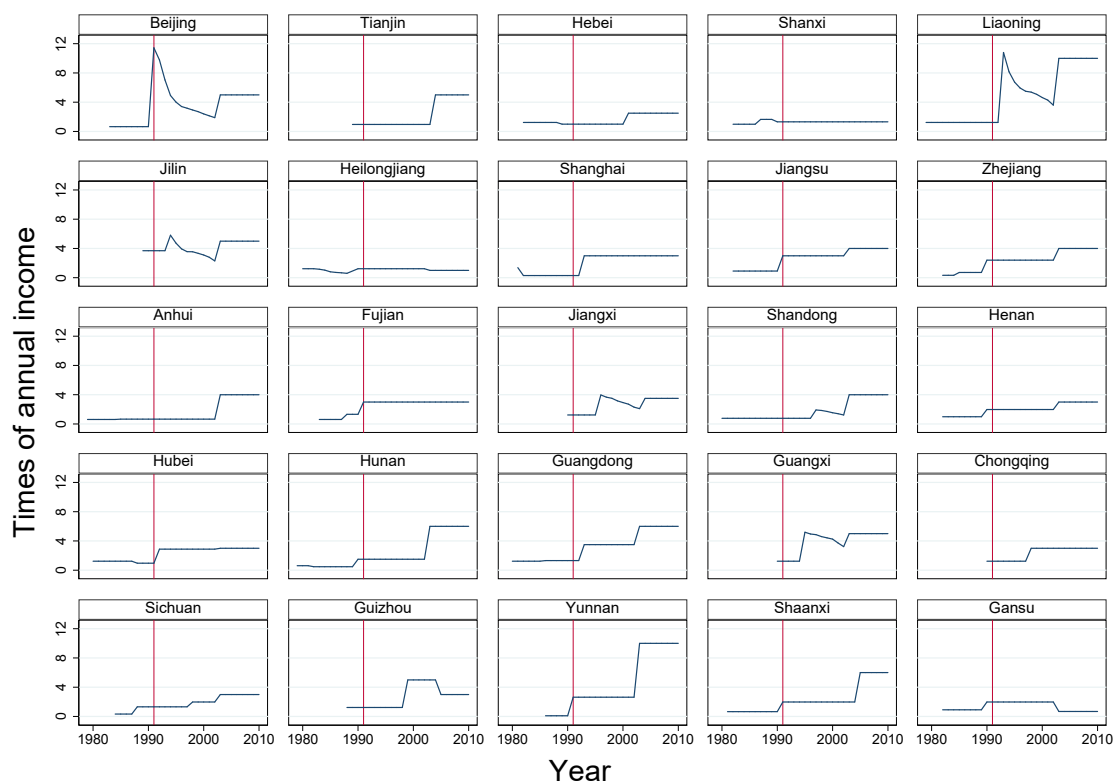
A third concern is that even though we find that there is more reduction in trust in local governments when the OCP was enforced more strictly in the first phase and there is no such effect in the second phase, the difference can be attributed to other factors rather than delegation. One potential factor is the performance of the government. It could be that the governments in the second phase performed better regarding economic development or public goods provision, and that is why people do not attribute responsibility to them. However, for this story to explain our results, it must be that the performance of the local governments is associated with the fertility penalty, which is our main measure of OCP exposure. As shown in Table 2.9, we did not find any significant correlation between the fertility penalty and measures of government performance, including GDP per capita, urbanization rate, and public expenditure.

2.7 Conclusion

This chapter provides field evidence on responsibility avoidance through delegation using evidence from China's one-child policy. Consistent with the predictions of the responsibility-shifting theory, we find that when local governments were the primary enforcer of the policy in Phase I, parents strongly exposed to the OCP currently trusted their local governments less. The OCP exposure did not undermine trust in neighbors in this phase as expected. However, when the local governments delegated the enforcement of the policy to civilians in Phase II, parents who were more exposed to the OCP currently trust their neighbors less. We do not find any significant impact of OCP exposure on trust in local governments among people who experienced the Phase II OCP. The differences in results between the two phases imply that responsibility avoidance through delegation is effective not only in the lab, as shown by previous studies, but also on a large-scale, highly influential field setting.

Authoritarian governments frequently mobilize the masses to enforce unpopular

policies. Civilians are encouraged and incentivized to report their neighbors, friends, and co-workers to the authority. For example, during Stalin's Great Purge, civilians were often sent to the Gulag as a result of reports initiated by friends and neighbors (Fitzpatrick, 1999). In Nazi Germany, homemakers, dentists, and other average citizens turned in their Jewish neighbors after petty neighborhood quarrels (Johnson, 2000). In a recent episode, the President of the Philippines, Rodrigo Duterte, encouraged vigilantes among the general population to commit violence against suspected drug users in his brutal drug war. In addition to grassroots information collection and shortage of police force, our findings provide another rationale for mass mobilization's popularity among the (authoritarian) governments - it helps the governments avoid part of the responsibility of implementing an unpopular policy.



Notes: The fertility penalties are measured in units of annual household income. For example, number 4 on the y-axis means a household needed to pay 4 times of its annual income after being caught violating the OCP. Data source: Scharping (2013) and authors' calculation.

Figure 2.1: Provincial Fertility Penalties in Urban China

2.8 Tables

Table 2.1: Summary Statistics

Variable	Obs	Mean	Sd
1991-2010			
Trust in neighbors ^a	1897	6.407	2.065
Trust in local governments ^a	1897	4.304	2.47
OCP exposure	1897	3.387	2.093
Age in 2016	1897	42.78	6.002
Family income per capita	1897	9.746	.867
Education attainment	1897	3.618	1.303
1979-1985			
Trust in neighbors ^a	822	6.582	2.072
Trust in local governments ^a	822	4.957	2.502
Birth control rate	822	89.164	3.101
Number of female at fertility age (000s)	822	78.33	39.37
Age in 2016	822	59.78	3.683
Family income per capita (in thousands)	822	27.71	34.79
Education attainment	822	2.94	1.168

a: categorical variables: 0 = extremely low trust;10 = extremely high trust

Table 2.2: Fertility Patterns in China

	Share of families having more than one child		Sex ratio of first birth
	Firstborn: boy	Firstborn: girl	
1990	0.479	0.538	1.052
2000	0.358	0.493	1.071
2005	0.285	0.418	1.024

Notes: Data is from China Census 1% sample (1990), 0.95% sample (2000), 1% sample (2005). We restrict our sample to married women who had their first child after 1979 and whose ages were between 21 and 40 in the census. We also restrict their matched children to those who aged between 0 and 18 in the census.

Table 2.3: Factors that Predict Gender of the First Child

Dependent variable	Having a first-born daughter		
	Probit	Probit	Linear prob.
	(1)	(2)	(3)
OCP exposure	0.024 (0.032)	0.017 (0.039)	0.007 (0.015)
Age		-0.002 (0.008)	-0.001 (0.003)
Log (income)		-0.029 (0.027)	-0.011 (0.010)
Years of education		0.010 (0.007)	0.004 (0.003)
Province FE	Y	Y	Y
Observations	1,726	1,726	1,726

Notes: *** p<0.01, ** p<0.05, * p<0.1. Table presents estimates of how observable characteristics predict the gender of the first child. The dependent variable is whether the first child is a girl. Robust standard errors are in parentheses.

Table 2.4: Estimates of OCP Exposure on Trust in Phase II
(1991-2010)

Dependent variable:	Trust in			
	Neighbors		Local governments	
	(1)	(2)	(3)	(4)
OCP exposure	0.045 (0.064)	0.033 (0.061)	-0.041 (0.080)	-0.053 (0.085)
Firstborn daughter × OCP exposure	-0.282*** (0.086)	-0.275*** (0.087)	-0.074 (0.156)	-0.073 (0.163)
p-Value	[0.003]	[0.004]	[0.639]	[0.656]
Bootstrap p-value	[0.004]	[0.004]	[0.684]	[0.724]
Individual controls		Y		Y
Cohort FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
Mean DV	6.407	6.407	4.304	4.304
Std.Dev.DV	2.065	2.065	2.470	2.470
Observations	1,897	1,897	1,897	1,897
R-squared	0.089	0.098	0.078	0.088

Notes: *** p<0.01, ** p<0.05, * p<0.1. OCP exposure in phase II is defined as the five-year mean value of the fertility penalty rates in province p after individual i had his/her first child. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. We use a wild cluster bootstrap-t procedure that are clustered at the province level for improved inference with a small number of clusters (Cameron, Gelbach and Miller, 2008). We report the corresponding p-values in brackets. We also report the p-values for OLS with clustered data. Number of clusters: 25.

Table 2.5: Estimates of OCP Exposure on Trust in Phase I
(1979-1985)

Dependent variable:	Trust in			
	Neighbors		Local governments	
	(1)	(2)	(3)	(4)
OCP exposure	0.100 (0.082)	0.086 (0.081)	0.109 (0.085)	0.098 (0.075)
Firstborn daughter × OCP exposure	-0.067 (0.087)	-0.037 (0.084)	-0.338*** (0.074)	-0.337*** (0.082)
p-Value	[0.449]	[0.667]	[0.000]	[0.000]
Bootstrap p-Value	[0.552]	[0.688]	[0.004]	[0.004]
Individual controls		Y		Y
Cohort FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
Mean DV	6.582	6.582	4.957	4.957
Std.Dev.DV	2.072	2.072	2.502	2.502
Observations	822	822	822	822
R-squared	0.142	0.155	0.140	0.156

Notes: *** p<0.01, ** p<0.05, * p<0.1. OCP exposure in phase I is defined as the five-year average family-planning rate in province p after an urban resident had his/her first child. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. We use a wild cluster bootstrap-t procedure that are clustered at the province level for improved inference with a small number of clusters (Cameron, Gelbach and Miller, 2008). We report the corresponding p-values in brackets. We also report the p-values for OLS with clustered data. Number of clusters: 25.

Table 2.6: Heterogeneous Effect of OCP Exposure on Trust in Phase II

Dependent variable:	Trust in			
	Neighbors		Local governments	
	One child	≥ Two children	One child	≥ Two children
	(1)	(2)	(3)	(4)
OCP exposure	0.039 (0.073)	0.006 (0.308)	-0.114 (0.085)	-0.206 (0.386)
Firstborn daughter ×OCP exposure	-0.258*** (0.079)	0.151 (0.384)	0.043 (0.173)	-0.306 (0.505)
Mean DV	6.370	6.558	4.227	4.621
Std.Dev. DV	2.067	2.058	2.454	2.509
Observations	1,516	380	1,516	380

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. OCP exposure in phase II is defined as the five-year mean value of the fertility penalty rates in province p after individual i had his/her first child. The sample in columns (1) and (3) is parents who have a single child. The sample in columns (2) and (4) is parents who have at least two children. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. Number of clusters: 25.

Table 2.7: Heterogeneous Effect of OCP Exposure on Trust in Phase I

Dependent variable:	Trust in			
	Neighbors		Local governments	
	One child (1)	≥ Two children (2)	One child (3)	≥ Two children (4)
OCP exposure	0.016 (0.107)	0.228** (0.106)	0.036 (0.057)	0.425 (0.268)
Firstborn daughter ×OCP exposure	0.001 (0.151)	0.072 (0.131)	-0.411*** (0.110)	-0.084 (0.323)
Mean DV	6.439	6.783	4.802	5.176
Std.Dev. DV	2.111	2.003	2.475	2.527
Observations	481	341	481	341

Notes: *** p<0.01, ** p<0.05, * p<0.1. OCP exposure in phase I is defined as the five-year average family-planning rate in province p after individual i had his/her first child. All regressions include a full set of province and cohort fixed effects. The sample in columns (1) and (3) is parents who have a single child. The sample in columns (2) and (4) is parents who have at least two children. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. Number of clusters: 25.

Table 2.8: Estimates of OCP Exposure on Trust in Central Government

Dependent variable:	Party membership	Media source for political news
	(1)	(2)
OCP exposure	0.004 (0.009)	-0.367*** (0.112)
Firstborn daughter × OCP exposure	0.007 (0.017)	0.147 (0.158)
Mean DV	0.112	-1.506
Std.Dev.DV	0.316	3.440
Observations	1,897	1,386
R-squared	0.195	0.122

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All regressions include a full set of province and cohort fixed effects. The dependent variable in column (1) is whether the individual is/was a member of the Communist Party. The dependent variable in column (2) is the difference between the number of days accessing political news through television and the number of days accessing political news through the Internet. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. Number of clusters: 19.

Table 2.9: Correlation between Fertility Penalty
and Government Performance

Dependent variable:	Fertility penalty
ln (GDP per capita)	0.409 (1.609)
Urbanization rate	-0.800 (0.994)
Share of public expenditure	1.170 (1.356)
Mean DV	2.600
Std.Dev. DV	2.071
Observations	534

Notes: *** p<0.01, ** p<0.05, * p<0.1. The regression includes year fixed effects. Urbanization rate is defined as the share of urban population. The share of public expenditure is defined as the share of government expenditure on culture, education, health, etc. Robust standard errors in parentheses. Number of clusters:25.

3.0 Correspondence Bias

When drawing inferences about a person's enduring characteristics from her actions, correspondence bias is the tendency to overly emphasize the role of the person's enduring characteristics and underestimate the influence of transient situational factors. Focusing on incentives as one important situational factor, we build a simple model to formalize correspondence bias, and we test predictions of the model in an online experiment. We randomly assign some subjects to play a 'benign' game that encourages cooperation and other subjects to play a 'malign' game that encourages selfish behavior. Everyone then needs to choose whether to take the dictator-game givings from an earlier stage of the experiment of a benign or malign game player. Consistent with correspondence bias, subjects are on average willing to pay to be matched with a benign-game player, an effect driven by both over-estimation of the prosociality of the benign-game player and under-estimation of the prosociality of the malign-game player. We show, further, that experiencing both games oneself, as opposed to only observing them, and receiving information about how each of the players behaved in both games, eliminates the effect.

3.1 Introduction

When drawing inferences about a person's enduring characteristics from their behaviors, the *correspondence bias* (Jones and Harris, 1967; Ross, 1977; Gilbert and Malone, 1995) is the tendency to overly emphasize the role of the person's characteristics and underestimate the impact of the incentives they face. For example, a common belief is that the rich are more selfish than the poor. They avoid taxes more often than others (Cox, 1984; Christian, 1994; Wang and Murnighan, 2014), and they break traffic laws more often when driving (Piff, Stancato, Côté, Mendoza-Denton and Keltner, 2012). However, the differences in behaviors can be fully explained by differences in incentives (Andreoni,

Nikiforakis and Stoop, 2017). The rich have more sources of income, which enables them to hide part of it, and paying traffic tickets is less painful for them due to diminishing marginal utility of money. The rich and the poor may actually be similarly generous, as shown beautifully by Andreoni, Nikiforakis and Stoop (2017), but people still attribute their selfish behaviors to their dispositions. Another good example is that voters often reward and punish incumbent presidents, governors, and senators for events that are completely beyond their control. Wolfers (2002) finds that voters in oil-producing states base voting decisions on the international price of oil, which should be beyond any governor's control.

Correspondence bias, which was previously overlooked by economists, has important implications for everyday life. Consider two regions (countries, neighborhoods, etc.) or ethnic groups where social norms (or institutions) are starkly different. Person A is from region/group 1 where unethical behaviors are harshly punished, and everyone finds it optimal to be trustworthy. Person B is from region/group 2 where legal enforcement is weak, and sabotaging others is common. Now a person C needs to choose one of the two to work with or hire. She observes that A behaved more reliably in the past than person B. If she is correspondence-biased, then she may jump to the conclusion that A is more trustworthy than B, and choose A over B even if B is a better match in other aspects.

We build a simple model to formalize the idea of correspondence bias. In our model, an individual chooses between two players after observing their actions, and the goal is to choose the one who is more likely to be the Good type. One player plays the *benign game* in which both the Good type and the Bad type choose to cooperate, while the other player plays the *malign game* in which the Good type cooperates and the Bad type defects. Borrowing the idea of *cursed equilibrium* (Eyster and Rabin, 2005), we model correspondence bias as the tendency to underestimate the correlations between actions and the game structure when interpreting the signals. Consequently, the biased individual tends to over-interpret cooperation in the benign game as a signal for the Good type and under-interpret cooperation in the malign game as a signal for the Good type. Even though the game one plays is completely determined by chance, we predict that a correspondence-biased individual is more likely to choose the benign-game player.

There are several challenges to empirically identify correspondence bias. Imagine an experiment in which we randomly assign half of the subjects to play the benign game that incentivizes everyone to cooperate, and the other half to play the malign game that incentivizes some people to defect. We then let them choose between a benign-game player and a malign-game player to play a follow-up game together. Our model predicts that on average people are more willing to pick the former. The first challenge is to distinguish between correspondence bias and Bayesian updating. Choosing the benign-game player can be consistent with Bayesian updating, as someone who chooses to cooperate in the benign game can rationally be expected to be more pro-social than someone who chooses to defect in the malign game. Second, reciprocity can also motivate choosing the benign game player. Participants may want to reciprocate the benign-game player's cooperative behavior in the follow-up game. Third, if one believes that people become more (less) pro-social under good (bad) institutions, as shown in (Peysakhovich and Rand, 2015; Cason, Lau and Mui, 2019), then it can make sense to choose the individual who played the benign game.

We deal with the three potential confounds using a three-stage experimental design. In the first stage, all subjects make a decision as the dictator in the dictator game. At the second stage, they are randomly matched into groups of four to play the benign game and the malign game. Both games are 2x2 complete-information games with a unique equilibrium in strictly dominant strategies. The malign game is the prisoners' dilemma game, while the benign game is the Harmony Game (Dal Bó, Dal Bó and Eyster, 2018) where the dominate strategy is to cooperate. At the end of the second stage, subjects are able to see the actions of a benign-game player and a malign-game player. In the third stage, they choose whether to be awarded the amount given in the first stage dictator game either from the benign-game player or the malign-game player, both of whose stage 2 actions are known to them. We use a multiple price list to elicit their willingness to pay for their preferred player.

We address the Bayesian updating confound by randomly assigning the players to the two games. The randomization ensures that the benign-game player and malign-game players are equally likely to be the Good type *ex ante*. The Martingale prop-

erty of Bayesian beliefs then implies that the expected posterior beliefs are the same. Therefore, a Bayesian model predicts that the individual will be indifferent between receiving the dictator offerings of the two players. However, our model predicts that a correspondence-biased individual is more likely to choose the benign-game player over the malign game player. Our design avoids the possibility that reciprocity could drive the results by using a dictator game in which there are no actions that the receiver can take; thus, there is no way to reciprocate the benign-game player's cooperation in the follow-up game. Finally, we avoid the potential for participation in one game or the other to shape the player's prosociality by sequencing the dictator decision so it occurs before stage 2. Even if individuals become more pro-social after playing the benign game, the dictator decision will have already been made in stage 1, and cannot be altered by the game.

To better understand correspondence bias, and to explore potential methods to reduce it, we utilize a 4-treatment design. The treatments differ in how many games one plays, and how much information one gets. In treatment 1, subjects only play one game, and are completely unaware of the other game. In stage 3, they are asked to choose between a benign/malign-game partner and a stranger. In treatment 2, subjects still only play one game as in treatment 1, but those who played the benign (malign) game also learn about the action of a malign-game (benign-game) player at the end of stage 2. In this treatment they are informed of the game played by the malign-game player, but they do not experience the game themselves. In stage 3 they choose between a benign-game player and a malign-game player given information about both players' actions in stage 2. In treatment 3, subjects play both games simultaneously in stage 2, and choose between their benign-game partner and malign-game partner in stage 3. Thus, not only do they experience the malign game themselves, but they actually play it with the malign-game player they have the option to choose in stage 3. The setup of Treatment 4 is the same as treatment 3, with the exception that subjects are also informed of both of their partners' actions in the other game (benign for the malign game partner and malign for the benign game partner) that each of their partners play with someone else.

Our results show, first, that correspondence bias exists and influences stage 3 de-

cisions. We measure the impact of correspondence bias through the *benign premium* – the amount a player, in stage 3, is willing to pay to receive the dictator game offering of the benign-game player, which they decided upon in stage 1. While the rational Bayesian model predicts the average benign-premium to be 0, we find that the benign premium is 12.48 cents on average in Treatment 2, the baseline treatment. It is significantly different from 0 at the 1% level. It means subjects are willing to pay 12.48 to get the benign-game player’s dictator transfers out of 100, the largest possible difference between the two potential dictators. Second, consistent with the predictions of our model, we find that correspondence bias is driven by both over-estimation of the prosociality of the benign-game player and under-estimation of the prosociality of the malign-game player. In treatment 1 when choosing between a stranger and a benign/malign-game player, subjects prefer the benign-game player to the stranger, and the stranger to the malign-game player.

We also test two potential methods to reduce or even eliminate correspondence bias. First, we study whether experiencing instead of observing the games can help to reduce the bias. Experiencing both games in Treatment 3, as opposed to only learning about it in Treatment 2, should help people to understand that actions are game-contingent, and they should take the games into account when inferring from the actions. Consistent with such an effect, we find that the benign premium in Treatment 3 is smaller than that in Treatment 2, although it is still significantly greater than 0 in Treatment 3 (at the 1% level), suggesting that experiencing both games is not enough to eliminate the bias. Second, we investigate the effect on reducing the bias of providing counterfactual information. In treatment 4, as subjects know both players’ action in both games, they should be able to compare their actions in the same games, which should alleviate the bias. We find that providing counterfactual information reduces the benign premium to 2 cents, which is not significantly different from 0, and is significantly smaller than that in Treatments 2 and 3.

The research that this study is most closely related to is Haggag, Pope, Bryant-Lees and Bos (2019)’s investigation of “Attribution Bias in Consumer Choice.” In their study, similar to ours, people underweight the impact of a transitory state on the utility of con-

suming a good, and misattribute it to the enduring characteristic of the good. The current research follows on their contribution by showing that attribution bias not only exists when it comes to evaluating consumption experiences, but also for evaluating people. Agents in Haggag, Pope, Bryant-Lees and Bos (2019)'s model do not fully appreciate the fact that their preferences are state-dependent; similarly, agents in our work fail to fully recognize that actions of other people are game-dependent. While Haggag, Pope, Bryant-Lees and Bos (2019) has important implications for individual decision making, our analysis shows that attribution bias also plays a vital role in economic interactions. We also add to their work by exploring potential debiasing methods. We show that experiencing both games instead of observing them reduces the attribution bias, and providing "counterfactual" information eliminates the bias. Our results provide an explanation for their finding that the extent of past experiences can attenuate the attribution bias in consumption choice.

Correspondence bias, previously designated the "fundamental attribution bias," has been intensively studied by psychologist since the 1960s (Jones and Harris, 1967; Ross, 1977; Gilbert and Malone, 1995). In a typical study, subjects listen to a speech arguing in favor of or against an opinion, and are asked to rate the attitudes of the speaker towards that opinion. A repeatedly-replicated finding is that, even when subjects are told that the speaker's positions are randomly assigned by the experimenter, they still rate speakers who are asked to speak in favor of the opinion as more supportive of it. The most common explanation for it in psychology is that the person (or internal characteristics) is more "salient" than the situation (or external factors), and thus people tend to underestimate the influence of the situation. We formulate the bias in a different way. We are less focused on the salience of other people's characteristic, but more on assessments of their stability. In our formulation in the following section of the paper, it is people's failure to fully account for the incentive-contingent nature of other's actions that leads them to under-attribute actions to situations.

The current study augments the existing psychology research on correspondence bias in three ways. First, the standard experimental paradigm for studying correspondence in psychology suffers from the potential confound that subjects may believe that

the randomly assigned positions can potentially shape the speakers' attitudes. As we discussed, our design rules this out. Second, in an environment that closely mimics real life interpersonal interactions, our design clearly shows that correspondence bias is welfare-reducing. Third, we provide a suggestion for how to reduce or eliminate correspondence bias by providing counterfactual information.

We also contribute to the literature on people's belief updating relative to Bayesian updating. Previous evidence suggests that people generally infer less from evidence than Bayes' Theorem predicts (Phillips and Edwards, 1966; Edwards, 1968; Möbius, Niederle, Niehaus and Rosenblat, 2014; Ambuehl and Li, 2018). However, as pointed out by Kahneman, this finding is in contrast to the everyday experience that people often jump to conclusions based on a little information. We provide another reason, in addition to the Law of Small Numbers (Kahneman and Tversky, 1972) and base-rate neglect (Kahneman and Tversky, 1973), for why people may draw overly extreme conclusions from small samples.¹ In our case, people jump from observations of other's action in narrow contexts to conclusions about those people's underlying qualities, without paying sufficient attention to the transient incentives they are facing.

The paper proceeds as follows: Section 3.2 describes a simple model of correspondence bias. Section 3.3 introduces our experimental design and the predictions it tests. Section 3.4 presents results, and Section 3.5 concludes and discusses policy implications.

3.2 Model

In this section, we build a simple model of correspondence bias that is in a similar spirit to the cursed equilibrium (Eyster and Rabin, 2005). In our model, the individual does not fully take into account the fact that other people's actions depend on the incentives they face (or the game they play). They are aware of the distributions of actions among others, but they underestimate the correlation between actions and the game structure when they try to interpret other people's actions.

¹For more discussion on over-inference, see Benjamin (2019).

Consider two games $\tau \in \{b, m\}$, the *benign game* b and the *malign game* m . In each complete information game, there are two actions to take: $\{C, D\}$. There are two types of agent, the good type G and the bad type B . Let the probability of being the good type be $P(t_i = G) = p_0$. In the benign game τ_b , both the good type and the bad type chooses C ; in the malign game τ_m , the good type chooses C and the bad type chooses D . Half of the players are assigned to play the benign game, and the other half are assigned to play the malign game.

After observing player i 's action in the benign game a_i^b and player j 's action in the malign game a_j^m , player k needs to choose between i and j to play a follow-up game. k 's payoff in the follow-up game is defined by the type of the partner of her choosing. If we normalize the payoff of choosing type B to be 0 and choosing type G to be 1, then player k 's payoff for choosing i to play the follow-up game is given by

$$U_{ki} = P(t_i = G) - P(t_j = G) \quad (3.1)$$

Without loss, the payoff for choosing j is normalized to 0.

Define $p(\cdot)$ as the true probability and $\pi(\cdot)$ as a person's potentially biased belief. For a Bayesian agent, the posterior can be calculated conditional on the games i and j played. As both types choose C in the benign game, the posterior $\pi(t_i = G \mid a_i^b) = p(t_i = G \mid a_i^b)$ is equal to the prior, p_0 . In the malign game, player j 's type is perfectly revealed. If she chooses C , then $\pi(t_j = G \mid a_j^m = C) = 1$; if she chooses D , then $\pi(t_j = G \mid a_j^m = D) = 0$. Therefore, $\pi_D^m < \pi_C^b < \pi_C^m$, where π_a^τ is the Bayesian agent's posterior belief about someone who chooses action a in game τ being the good type.

The correspondence-biased agent knows the distribution of actions across the whole population, but they cannot fully account their opponent's action based on the game they played. Borrowing from Eyster and Rabin (2005)'s *cursed equilibrium*, we define an agent as χ -biased if

$$\tilde{\pi}_a^\tau = \chi\left(\frac{1}{2}\pi_a^b + \frac{1}{2}\pi_a^m\right) + (1 - \chi)(\pi_a^\tau) \quad (3.2)$$

where $\tilde{\pi}_a^\tau$ is the correspondence-biased agent's posterior belief about the player who chooses action a in game τ , $\frac{1}{2}$ is the probability one is assigned to play the benign game or

the malign game. Intuitively, χ -biased agent only recognizes the action of her opponent but ignores the incentive/game structure she faces with probability χ . In this case, she replaces the actual probability $p_a^\tau = \pi_a^\tau$ of her opponent being the good type given action a in game τ with the average probability $\frac{1}{2}\pi_a^b + \frac{1}{2}\pi_a^m$ of her opponent being the good type given action a across the two games. If $\chi = 0$, then the χ -biased agent's posterior is the same with the Bayesian agent. As long as $\chi > 0$, we can conclude that

$$\pi_D^m = \tilde{\pi}_D^m < \pi_C^b < \tilde{\pi}_C^b \leq \tilde{\pi}_C^m < \pi_C^m \quad (3.3)$$

, where $\tilde{\pi}_C^b = \tilde{\pi}_C^m$ when $\chi = 1$.

Therefore, the probability of the benign game player i being the good type given she chooses C is overestimated, or $\pi_C^b < \tilde{\pi}_C^b$. The probability of the malign game player j being the good type given he chooses C is underestimated, or $\tilde{\pi}_C^m < \pi_C^m$. It is important to note that the χ -biased agent can still make the inference that a_C^m is a stronger signal for the good type than a_C^b , i.e. $\tilde{\pi}_C^b \leq \tilde{\pi}_C^m$, but they underestimate the difference between the two: $\tilde{\pi}_C^m - \tilde{\pi}_C^b < \pi_C^m - \pi_C^b$.

The correspondence bias implies an over-evaluation of the payoff for choosing the benign-game player. Depending on the malign-game player j 's choice, there are two cases: $a_i^b = C, a_j^m = D$ and $a_i^b = C, a_j^m = C$. In the first case, the Bayesian agent should conclude that the payoff for choosing the benign-game player equals to

$$U_{ki} = \pi_C^b - \pi_D^m \quad (3.4)$$

However, the correspondence biased agent would believe that the payoff of choosing i , \hat{U}_{ki} , is

$$\hat{U}_{ki} = \tilde{\pi}_C^b - \tilde{\pi}_D^m \quad (3.5)$$

As $\tilde{\pi}_C^b > \pi_C^b$ and $\tilde{\pi}_D^m = \pi_D^m$, $\hat{U}_{ki} > U_{ki}$. Similarly, when $a_i^b = C, a_j^m = C$, as $\tilde{\pi}_C^b > \pi_C^b$ and $\tilde{\pi}_C^m < \pi_C^m$, \hat{U}_{ki} is also larger than U_{ki} .

For a Bayesian agent, the expected benefit of choosing i should be 0. The Martingale

property of Bayesian updating implies that

$$E[\pi \mid \tau = b] = E[\pi \mid \tau = m] = p_0 \quad (3.6)$$

Intuitively, as whether one plays the benign game or the malign game is completely determined by chance, i and j are equally likely to be the good type *ex ante*. As expected posterior is equal to the prior, they are equally likely to be the good type in expectation *ex post*.

However, for a correspondence-biased agent, the expected benefit of choosing i is larger than 0. As $\tilde{\pi}_C^b > \pi_C^b = p_0$ for the biased agent, $E[\tilde{\pi} \mid \tau = b] > p_0$. As $\pi_D^m = \tilde{\pi}_D^m$ and $\tilde{\pi}_C^m < \pi_C^m$, $E[\tilde{\pi} \mid \tau = m] < E[\pi \mid \tau = m] = p_0$. Therefore, for the correspondence-biased agent

$$E[\tilde{\pi} \mid \tau = m] < p_0 < E[\tilde{\pi} \mid \tau = b] \quad (3.7)$$

We summarize our main results in the following proposition.

Proposition 1. i) For any $\chi \in (0, 1]$, a χ -biased individual is willing to pay an expected benign premium of $E[\tilde{\pi} \mid \tau = b] - E[\tilde{\pi} \mid \tau = m] > 0$ to be matched with the benign-game player.

ii) For any $\chi \in (0, 1]$, a χ -biased individual is willing to pay an expected premium for the benign-game player when choosing between her and a stranger, and is willing to pay an expected premium for the stranger when choosing between him and the malign-game player.

3.3 Design

The experiment has three stages. At the first stage, all subjects make a decision as the dictator in the dictator game. At the second stage, they are randomly matched into groups of 4 to play the benign game and the malign game. They are encouraged to cooperate with their partners in the benign game, but are motivated to behave selfishly in the malign game. Lastly, they are asked, as the receiver, to choose a dictator between a benign game player and a malign game player to play the first stage dictator game. Our

model predicts that there exists a *benign premium*: subjects are on average willing to pay to be matched with the benign-game player.

3.3.1 First Stage

The experiment is conducted online, and subjects are recruited through Amazon Mechanical Turk (Mturk). Upon arriving at the study website, subjects are instructed to play a dictator game as the dictator. They need to divide 200 cents between themselves and a random receiver. As in a standard dictator game, the receiver has no influence over the outcome of the game, and both the receiver and the dictator receives 50 cents of endowment prior to the split decision. Subjects are also informed that even though everyone needs to make the decision, only half of those decisions will be implemented later. At this stage, they have no idea of the future stages or the identity of the potential random receiver. This dictator decision serves as our measure of one's prosociality.

3.3.2 Second Stage

At the second stage, subjects are randomly matched into 4-player groups. Everyone is randomly assigned a role. There are 4 roles in each 4-player group. We name them A, B, C and D. Then the participants play the benign and/or the malign games with members in their own group. Depending on the treatment, a subject may interact with 1 or 2 roles at this stage. One's own role and the roles one play with in this stage are common knowledge. The two games are defined as follows.

The malign game is a two-player one-shot Prisoner's Dilemma game—see left panel in Table 3.1. Participants in this game must choose between cooperate (C) and defect (D).² It is a dominant strategy to choose D for both players. The Nash equilibrium of this game is (D,D), which leads to the payoffs (30,30). Even though D is the dominant strategy in this game, previous studies show that not everyone chooses to defect, and those who choose to cooperate are more prosocial than those who choose to defect (Cooper,

²The actions C and D were respectively labeled "Action 1" and "Action 2" in the experiment to ensure a neutral presentation.

DeJong, Forsythe and Ross, 1996; Barreda-Tarrazona, Jaramillo-Gutiérrez, Pavan and Sabater-Grande, 2017). Conveniently, we define subjects who choose to cooperate in the malign game as the *Good* type and subjects who choose to defect in it as the *Bad* type.

The benign game is the Harmony game as in Dal Bó, Dal Bó and Eyster (2018) - see right panel in Table 3.1. Participants also choose between cooperate (C) and defect (D). However, (C,C) is the dominant strategy equilibrium and Pareto dominates all other strategy profiles in this game. It is easy for the subjects to figure out that they should choose C.

To ease understanding, we will illustrate the rest of the experimental design in terms of Treatment 2, which we believe to be closest to reality. We will then discuss the differences between the treatments at the end of this section. In treatment 2, subjects play one game, and observe outcomes of the other game. To be more specific, player A and B play the benign game, and C and D play the malign game. Even though they only play one game, we also give them the instructions of the other game. Therefore, they can still understand the incentives of the game they do not play.

After the actions being taken, subjects enter the information provision page. In this page, they learn about their payoffs in the game they play and the action of their partner in that game. We also display the payoff table of the two games again to minimize confusions and incorrect recalls. In addition, they are also informed of the action of one of the two players in the game they do not play. To be more precise, A (B) in the benign game also learns whether D (C) in the malign game chooses to cooperate or defect. Similarly, C (D) also learns the action of B (A) in the benign game.

Subjects are forced to stay at least 120 seconds on the information provision page to make sure that they take the time to understand the game structure, and make inferences about the types of their partners based on their actions.

3.3.3 Third Stage

At the third stage, every subject has the opportunity to choose a dictator's transfers in the first stage between a malign-game player and a benign-game player. The two candi-

dates are those two whose actions in stage 2 were shown to the subject in the information provision phase. For example, A played the benign game with B and observed D's action in the malign game. Then in stage 3, A needs to choose between B and D's dictator transfers in the first stage dictator game. Therefore, when making the choice between the two candidates, the subject can rely on the inferences she made from their actions in stage 2.

After reading the instructions for stage 3 and before making any decisions, subjects are asked to answer 3 comprehension questions. Only those who answer all the 3 questions correctly can proceed to make their choices in this stage. Those who at least answer one question incorrectly need to re-do all the 3 questions again until they get them all correct.³

After a subject makes the choice between the two candidates, we use a multiple price list to elicit her willingness to pay to be matched with the partner of her choosing. The list shown to A if she chooses B over D in the first choice is displayed in Table 3.2 as an example. In total, subjects need to make 10 additional choices after the first choice. In each choice, there are two options. $D+(x \text{ cents})$ means if in this choice A chooses D, then she can get an extra reward of x cents. But if she chooses B, there is no extra reward. The point where A switches from option 1 to option 2 defines A's willingness to pay to get B. For example, if A chooses option 1 for choice 1 to 7 and option 2 for choice 8 to 10, then it means A believes that B gave 70 to 80 more than D in the dictator game on expectation. One of these 11 choices (including the one between B and D with no extra rewards) will be randomly selected as the *choice-that-counts*.

As every member of the group makes a choice between 2 potential dictators, there are 4 choice-that-counts in total. We randomly select 1 of the 4 to implement. If one's choice-that-counts is chosen, then she gets the dictator's transfers of her choosing plus the extra rewards. The candidate she did not choose also becomes the dictator, and the remaining player of the 4 becomes another receiver. For example, if A's choice-that-counts is

³Please see Appendix xxx for the comprehension questions. As one cannot proceed to the decision stage of stage 3 without answering all 3 questions correct, some subjects dropped out in this stage. Out of 1037 subjects who signed up for the experiment, 158 of them only finished stage 2. As stage 3 is not interactive, the dropouts of those subjects have no impact on the rest in the same groups.

Choice 5, she chooses Option 2 and her choice-that-count is chosen to be implemented, then A will be matched with D with a 60 cents extra reward. The two remaining participants are automatically matched afterwards with B being the dictator as B was one of A's candidates.

The dictator's decisions which were made at the first stage are then carried out. For example, suppose B chose to give x cents to the random receiver at the first stage, then A gets x cents and B gets $(200 - x)$ cents if A and B are matched with B being the dictator.

3.3.4 Three Confounders

Our three-stage design deals with three challenges to identify the correspondence bias. As predicted by our model, correspondence bias leads to a benign premium: a positive willingness to pay to get the benign-game player instead of the malign-game player. However, other forces can also produce a benign premium. The first possibility is choosing the benign-game player can be consistent with Bayesian updating as someone who chooses to cooperate in the benign game is expected to be more pro-social than someone who chooses to defect in the malign game. Second, reciprocity can also motivate choosing the benign game player. Participants may want to reciprocate the benign-game player in the follow-up game as he/she has been nice to them (or another player) in the benign game. Similarly, they may also want to punish the malign-game player for choosing the selfish option. Third, institution can also shape people's prosociality (Peysakhovich and Rand, 2015; Cason, Lau and Mui, 2019). Even if the benign-game player and the malign-game player are *ex ante* the same, the benign-game player can be shaped into a nicer person by the game, and it makes sense to choose her over the malign-game player.

As we only look at people's average willingness to pay to be matched with the benign-game player and subjects are randomly assigned to play the benign game or the malign game, Bayesian updating cannot explain a positive benign premium. The intuition is as subjects are randomly assigned to the games, the two players are equally likely to be the Good type *ex ante*. According to the martingale property of Bayesian belief, expected

posterior equals to prior, which implies that the two players are equally likely to be the Good type on expectation *ex post*. Therefore, on expectation, Bayesian subjects should treat the benign-game player and the malign-game player equally.

The feature that there is no choice for the receiver to make in the dictator game helps us deal with the reciprocity confounder. In the third stage when choosing the dictator between the two candidates, even if the subject want to reciprocate the benign-game player there is no way for them to do so. First, as there is no choice in the dictator game for the receiver, she cannot choose the benign-game player and then be nice to her in the follow-up game. Second, one may worry that choosing the benign-game player as the dictator in the third stage *per se* can be seen as reciprocity. Even though a player is indeed expected to earn more being the dictator than being the receiver, choosing him/her in the third stage does not increase his/her chance of being the dictator. As illustrated earlier, which two players become the dictators is solely determined by chance. One's choice in the third stage only affects which one of the two becomes her own dictator (in case he/she is chosen to be one of the two dictators).

Putting the dictator decision ahead of the second stage games solves the third confounder that institution shapes one's prosociality. The idea is at the third stage, the dictators have already made their decisions about how much to transfer in the first stage. Therefore, what happens at the second stage cannot have an impact on them. Even if the benign-game player becomes a nicer person after playing the game, her choice in the first stage remains the same. All the subjects have to do is to draw inferences about the players' types from the second stage, and choose the one who they believe to be more pro-social in Stage 1.

3.3.5 Treatments

There are four treatments in the experiment and they only differ in the second stage. What separate them from each other is how many games each subject plays and how much information they get.

As stated earlier, in Treatment 2 each player only plays one game. In the information

stage, in addition to the outcomes and the action of the opponent in the game one plays, he/she also observes the action of one other player who plays the other game.

In Treatment 1 (as indicated in Figure 3.1), similar to Treatment 2, each player only plays one game, either the benign or the malign game. However, in this treatment subjects are not aware of the existence of the other game. At the third stage, subjects will be asked to choose between their opponent in the second stage and a random stranger.

In Treatment 3, all the settings are the same with Treatment 2 except that subjects play both games in the second stage. In the information stage, they learn the actions of both their opponents and their outcomes in the two games, which is equivalent to Treatment 2 except that all information are gathered through “experience” instead of partly through “observation” as in Treatment 2.

In Treatment 4, all settings remain the same with one exception: subjects are also informed of the behaviors of their opponents in the game they does not play together. In Treatment 3, A (C) plays the benign game with B (D) and the malign game with D (B). But A (C) cannot observe which action B (D) chooses in the malign (benign) game with C (A), or which action D (B) chooses in the benign (malign) game with C (A). While in Treatment 4, all such information is available. Therefore, when A chooses between B and D, A not only knows the action they choose when they are playing with her, but also is aware of their behaviors in the alternative situation.

3.3.6 Predictions

The four-treatment design helps us investigate the causes of the correspondence bias and the potential ways to reduce or even eliminate it.

Treatment 2 is our baseline treatment, and we can test the existence of correspondence bias by looking at the benign premium in this treatment.

Prediction 1. There exists a benign premium in Treatment 2, i.e., the average willingness to pay towards the benign-game player is larger than 0.

Treatment 1 aims to decompose the benign premium. As no information is provided on the stranger, the chance of her being the Good type equals to the prior, p_0 . Thus,

Treatment 1 helps us separate benign premium $E[\tilde{\pi} | \tau = b] - E[\tilde{\pi} | \tau = m]$ into two parts: under-estimate of the chance of the malign-game player being the Good type $p_0 - E[\tilde{\pi} | \tau = m]$ and over-estimate of the chance of the benign-game player being the Good type $E[\tilde{\pi} | \tau = b] - p_0$. While Bayesian inference predicts that $E[\pi | \tau = m] = E[\pi | \tau = b] = p_0$, we predict that for correspondence-biased agents $E[\tilde{\pi} | \tau = m] < p_0 < E[\tilde{\pi} | \tau = b]$.

Prediction 2. In Treatment 1, the average willingness to pay for the benign-game player when choosing between her and a random stranger is positive; the average willingness to pay for the malign-game player when choosing between him and a random stranger is negative.

Treatment 3 is set to test whether inattention to strategic motives is a cause of correspondence bias. As participants play both games in this treatment, they have a better understanding of the incentives in the two games. In Treatment 2, the subject may only pay attention to behaviors without understanding the incentives behind them. Consequently, she tends to treat cooperation in the two games equally even though it is a much stronger signal of the Good type to cooperate in the malign game. When she play both games herself in treatment 3, she is more likely to know that choosing cooperation does not mean the same thing across the two games. We therefore predict that subjects who only play the malign game in treatment 2 pay the same benign premium as the average subject in treatment 3; while subjects who only play the benign game in treatment 2 are willing to pay a higher benign premium than an average subject in treatment 3.

Prediction 3. The benign premium is smaller in Treatment 3 than in Treatment 2.

In Treatment 4, we test whether providing counter-factual information reduces the correspondence bias. In Treatment 2 and 3, the participant is not able to know how the benign-game player performs in the malign game, and *vice versa*. However, in treatment 4, such information is available, and subjects can clearly see how other's actions change according to the incentives. If the correspondence bias is caused by failing to fully account for the impact of the incentives on actions, then enabling people to compare opponents' behaviors in the same game with the same incentives should reduce the bias significantly.

Prediction 4. The benign premium is smaller in Treatment 4 than in Treatment 3.

3.4 Results

The experiment was conducted on Amazon Mechanical Turk between October 12th 2018 and December 7th 2018. We also included data from two earlier pilots on June 11th 2018 and July 17th 2018. As our experiment is rather complicated, we only recruit subjects who at least have a two-year associate degree. We also restrict our participants to those who are residents in the United States. On average, the experiment lasts 20.1 minutes and subjects earn 2.77 dollars on average. Overall, we recruited 839 subjects, with 133 in treatment 1, 251 in treatment 2, 228 in treatment 3 and 227 in treatment 4.⁴⁵

For a manipulation check, we look at the cooperation rate in the two games. While almost everyone chose to cooperate in the benign game (94.87%), the frequency of choosing cooperation was much lower in the malign game (41.05%). Subjects who chose cooperation in the malign game is indeed more pro-social than those who chose to defect. As shown in Figure 3.4, subjects in the benign game on average transferred 66.18 cents in the first stage. While subjects who chose C in the malign game transferred 76.51 cents in the first stage, subjects who chose D only transferred 59.04 cents. The average transfers by the malign-game players were 66.21 cents, which is very close to the benign-game player's average.

Table 3.4 shows the summary statistics and the balance check. It indicates that there's no significant difference in the amount of money people choose to keep in the dictator game in stage one among the treatments. A nature concern is that subjects may change their behaviors between treatment 2 and treatment 3 due to the number of games they play. The cooperation rates in the malign game in the two treatments are not significantly different from each other (p-value=0.753).

Result 1: Correspondence bias exists in the baseline treatment. We first look at the existence of correspondence bias in our baseline treatment, Treatment 2. As predicted

⁴⁵We received a total of 879 responses, but dropped 40 subjects (4.6%) who exhibited multiple switching points in the multiple price-list questions at the third stage.

⁵We randomly assign less subjects to Treatment 1 based on a power calculation. We need more subjects in the other 3 treatments because we need to test whether the benign premium is significantly different between two treatments. While in Treatment 1, we only need to test whether the average willingness to pay is significantly different from 0 or not.

by a rational Bayesian model, subjects should be indifferent between the benign-game player and the malign-game player on average when choosing their dictator in stage 3. However, our model predicts a benign premium: the average willingness to pay towards the benign-game player is larger than 0. Using the multiple price list, we define the benign premium as the switch point between Option 1 and Option 2 in Table 3.2. We further code the benign premium as positive if a subject chooses the benign-game player in the first choice, and negative otherwise. Since the multiple price list can only elicit intervals of benign premium, our preferred measure uses the mid-point of the interval as the benign premium.⁶ For example, if subject A chooses B (the benign-game player)'s transfer over D (the malign-game player)'s transfer plus 10 cents bonus, and she switches to D's transfer plus 20 cents, then A's benign premium is coded as 15 cents.

As shown in Figure 3.2, the average willingness to pay towards the benign-game player is 12.48 cents in Treatment 2, which is significantly larger than 0 at the 1% level. One way to interpret this result is subjects on average believe that the benign-game player transferred 12.48 more cents in stage 1 than the malign-game player. To put those numbers into perspectives, one can compare them with the maximum possible benign premium 100 cents.⁷

We further explore the share of subjects who are correspondence-biased in treatment 2. It is hard to tell someone is biased or not when the malign-game player chooses D. Both the Bayesian model and our model predict that subjects should choose the benign-game player, and the only difference is our model predicts a larger willingness to pay towards the benign-game player. However, the case when the malign-game player choose C is more clear-cut. While a Bayesian subject should choose the malign-game player for sure regardless of her prior, our model predicts that a fully correspondence-biased subject is indifferent between the two players and may choose the benign-game player. Our data shows that there are still 52.88 % of subjects choose the benign-game player over the malign-game player when the latter chooses to cooperate.

⁶Our results are robust if we use the minimum (maximum) value of the interval as the benign premium (Figure 3.5).

⁷A completely selfish individual transfers 0 in stage 1, while an altruistic individual who weights other's utility exactly as much as her own transfers 100 cents in stage 1. Therefore, the largest possible difference between the two potential opponents' transfer is 100 cents.

Result 2: Evidence suggests that the correspondence bias is caused by subjects' failures to fully account for the correlation between actions and incentives. In Treatment 1, subjects only play one game, and are asked to choose between their partner and a random stranger in stage 3. As predicted by the model, a Bayesian subject should be indifferent between her partner and the stranger on average regardless which game they played together. However, the game one is assigned to play does have an impact on one's willingness to pay towards the stranger and the WTP is not 0.

Treatment 1 is more comparable to previous studies in psychology on the correspondence bias. We randomly assign the subjects to interact with someone in a benign environment or a malign environment, and we aim to test whether this randomly assigned environment has an impact on subject's evaluation on their partner or not. Our results show that the orthogonal environment has a strong effect on subjects' WTP towards their partner. When the game played together is the benign-game, the average willingness to pay for the partner over the stranger is 13.82; when the game is the malign-game, the average WTP for the partner is -8.85, meaning that subjects are willing to pay to get the stranger, instead of their partner, as their dictator. The two WTPs are significantly different from each other ($p\text{-value} < 0.001$, Wilcoxon ranksum test), which serves as another piece of evidence for the correspondence bias.

Treatment 1 also serves as a test of our formulation of the correspondence bias. If the bias is caused by people's failure to fully account for the degrees to which incentives affect actions, then we would predict that they prefer the benign-game player to the stranger and prefer the stranger to the malign-game player. Our results are consistent with this prediction. As showed above, the WTP for the benign-game player is positive and is significantly different from 0, with a $p\text{-value}$ of 0.009. Meanwhile, the WTP for the malign-game player is negative and is also significantly different from 0 ($p\text{-value} = 0.072$). The negative WTP for the malign-game player is unlikely to be a mistake as subjects do respond to the malign-game player's actions. When the malign-game player chooses to cooperate, the average WTP towards her is 12.6; when the malign-game player chooses D, the average WTP is -22.25.

Result 3: Experience reduces the correspondence bias, but it alone is not able to

eliminate the bias. Previous results jointly show that there exists a correspondence bias: subjects tend to believe that someone who is randomly assigned to play a benign game is more pro-social than someone who is randomly assigned to play a malign game. The next question is can we alleviate this bias? By comparing treatment 2 with treatment 3, we can see the effect of letting the subjects experience both regimes to understand the strategic motives better. The only difference between the two treatments is that subjects only play one game but observe the other one in treatment 2, while in treatment 3 they play both. The average benign premium decreases from 11.57 in Treatment 2 to 7.98 in Treatment 3, with a p-value of 0.298.

The reduction in benign premium from Treatment 2 to 3 is mainly driven by the reduction in benign premiums of subjects whose malign-game player chooses to defect. As shown in Figure 3.3, when the malign-game player chooses D, the benign premium reduces from 20.21 to 15.23 (p-value=0.195). Meanwhile, the benign-premium only reduces from -0.05 to -1.98 when the malign-game player chooses to cooperate. The results suggest that experience is better at reducing the over-estimation of the niceness of the benign-game player. It has little effect on reducing the under-estimation of the niceness of the malign-game player.

One potential concern is the difference between treatment 2 and treatment 3 can also be driven by inattention to partner's choices instead of inattention to the strategic motives. Subjects might only pay attention to the game they played and ignored the other game. To deal with this issue, we further look at how subjects who only played the benign-game respond to malign-game players' choices. If they are paying no attention to the malign-game player's choices, then the benign premium should be the same regardless of the choices. Table 3.5 shows that for benign-game only players, when the malign game player choose to cooperate, their benign premium is 13.89; while when they choose to defect, the benign premium increases to 26.84. The two amounts are significantly different from each other (p-value=0.065).

Experience alone is not sufficient to eliminate the correspondence bias. The benign premium in Treatment 3 is still significantly larger than 0 (p-value=0.002, t-test).

Result 4: Providing counterfactual information in addition to letting subjects ex-

perience both games can eliminate the correspondence bias. The result is mainly driven by reduction in over-estimation of the niceness of the benign-game player. By comparing treatment 3 and treatment 4, we can study the effect of informing the subjects the “counterfactual”. When we not only let subjects learn the behaviors of two partners by playing games with them as in treatment 3, but also inform them the behaviors of their partners in the game they did not play together in treatment 4, the benign premium further decreases to 2.14, which is not significantly different from zero ($p=0.401$). The difference between Treatment 3 and Treatment 4 in benign premium is significant at the 10% level ($p\text{-value}=0.077$), suggesting providing counterfactual can alleviate the correspondence bias. The difference between Treatment 2 and Treatment 4 is significant at the 1% level ($p\text{-value}=0.007$), which indicates that experience plus counterfactual information can jointly eliminate the bias.

The reduction in WTP for the benign-game player between Treatment 3 and 4 is mainly driven by the reduction in benign premiums of subjects whose malign-game player chooses to defect. The WTP reduces from 15.23 to 6.37, and the difference is significant at the 10% level ($p\text{-value}=0.093$). The WTP for the benign-game player in Treatment 4 is very close to the rational level with the correct prior. As 3.4 illustrates, the difference in stage 1 transfer between the benign-game player (66.18) and the malign game player (59.04) who choose to defect is 7.14. It shows that the over-estimation of the niceness of the benign-game player is almost gone. At the same time, when the malign-game player chooses to cooperate, the benign premium reduces from -1.98 in Treatment 3 to -3.65.⁸

This result that providing the “counterfactuals” help to de-bias also explains why we always observe the correspondence bias in real life: it is impossible to observe the “counterfactual” in everyday life. For example, in a society with low mobility, the rich is born rich and the poor may remain poor. It is hard to see how the rich would behave if they were poor, and it is hard to observe how the poor would behave if they got rich. Even if some people experienced both cases, it is hard for others to witness how they behaved in the two different situations.

⁸Again, it is also closer to the rational amount with correct prior, $-10.33 = 66.18 - 76.51$.

3.5 Conclusion

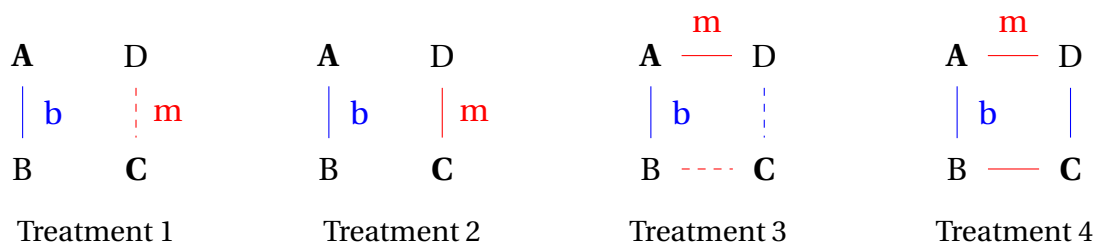
This chapter investigate people's tendency to underestimate the correlation between other's actions and the incentives he/she faces when drawing inferences about other's enduring characteristics from their actions. We study this bias in an environment in which some people are assigned to play a benign game that encourages cooperation and other people are assigned to play a malign game that encourages selfish behaviors. While a rational Bayesian model predicts that people should pick the benign-game player and the malign-game player with equal chances in a follow-up game, our model predicts that a correspondence-biased individual chooses the one who played the benign game more often in the follow-up game. The key intuition is failing to fully appreciate the impact of the incentives on actions leads the biased individual to over-interpret cooperation in the benign game and under-interpret cooperation in the malign game.

We test the predictions of the model in an experiment with 852 subject. We first ask subjects to make a decision about how much to transfer as the dictator in a dictator game. Then we let them play the benign game and the malign game, and inform them the actions of a benign-game player and a malign-game. Lastly, we ask them to choose, as a receiver in the dictator game, between the benign-game player and the malign-game player's first stage transfer. We find strong evidence for correspondence bias. One is willing to pay 12.48 out of 100 to be matched with the benign-game player in Treatment 2, our baseline treatment. Allowing subjects to experience both games instead of playing one and observing the other one reduces the correspondence bias, but the benign premium is still significantly above 0. However, if we inform subjects how their benign-game partner behaves in the malign-game and *vice versa*, the correspondence bias disappears.

Both Treatment 3 and 4 suggests that ignorant to the effect of incentives on actions is the cause of the correspondence bias. In Treatment 1, we directly test the predictions of our model. We find that subjects are willing to pay to be matched with the benign-game player when choosing between her and a stranger, and they are willing to pay to be matched with the stranger when choosing between him and the malign-game player. This finding is hard to be reconciled with other theories of the correspondence bias.

Our findings shed lights on both why the correspondence bias is widely observed in real life and the potential ways to reduce or eliminate it. First, in reality, we often only experience one environment and observe other environments, which makes it hard for us to understand how alternative environments affect other people's behaviors. Accordingly, one potential route to speed up social cohesion is to encourage social interactions between different groups and let them experience other people's lives (Rao, 2019; Lowe, 2020). Second, counterfactual information is hard to get in reality. We often only interact with others in one situation without knowing how they perform in completely different cases.

3.6 Figures

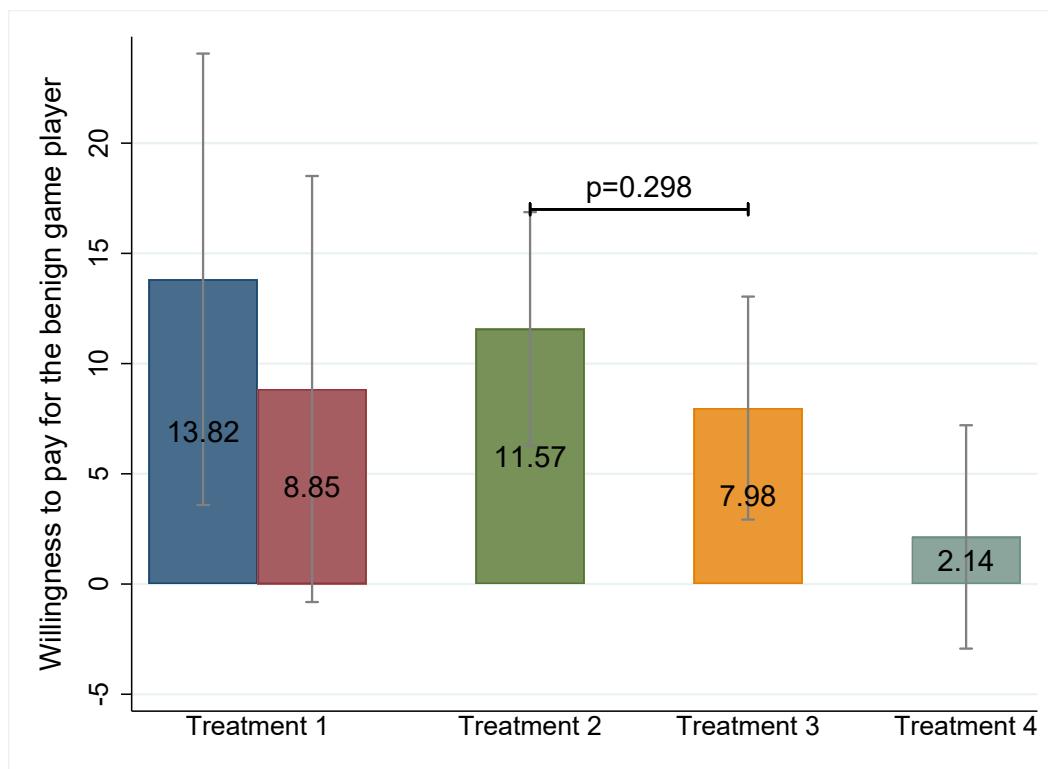


b: benign game **m:** malign game

— : observed by A - - - : not observed by A

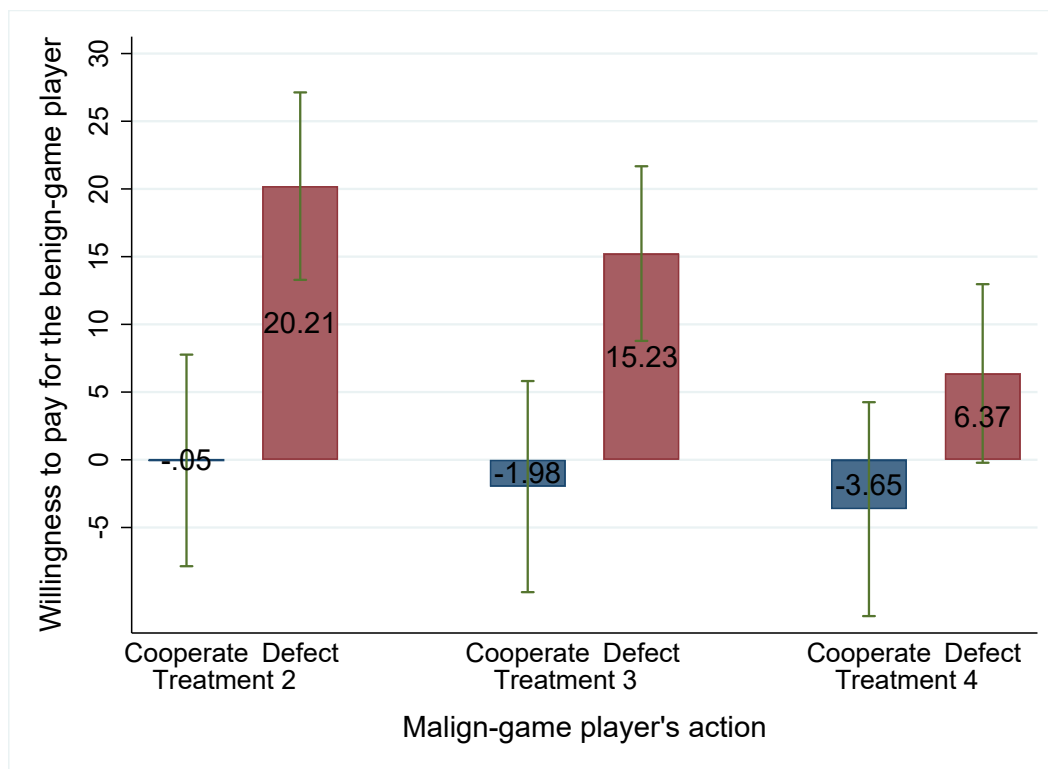
Notes: Figure displays the four treatments in subject A's perspective. The solid line means that A is able to observe (the outcome of) a game and the dash line means that A is not able to observe a game. But of course, A is not the only active player in the game. The game faces by B, C and D are symmetric in treatment 2, 3 and 4. For example, player D plays the benign game with C and the malign game with A in treatment 3 and she cannot observe the game played between A and B or the game played between C and B. The game is not symmetric in treatment 1. In that treatment, A and B only play the benign game and C and D only play the malign game.

Figure 3.1: Four Treatments



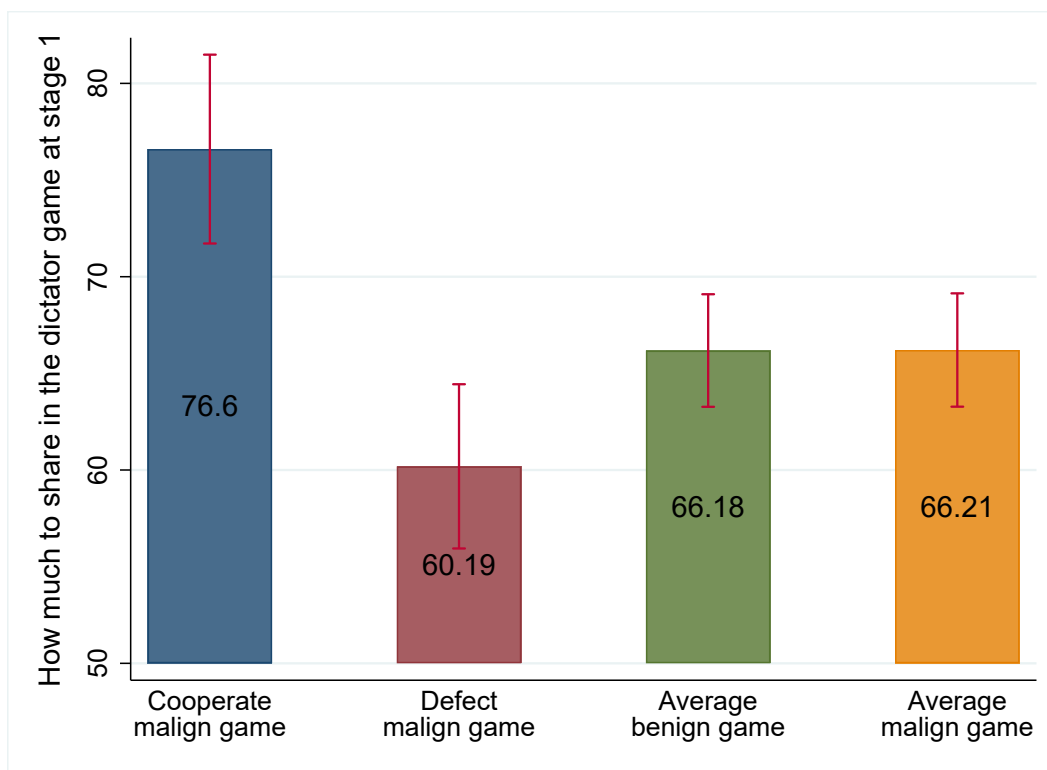
Notes: The left bar in treatment 1 represents the WTP for the benign game player versus a stranger, and the right bar in treatment 1 represents the WTP for a stranger versus a malign game player.

Figure 3.2: Benign Premiums across Treatments



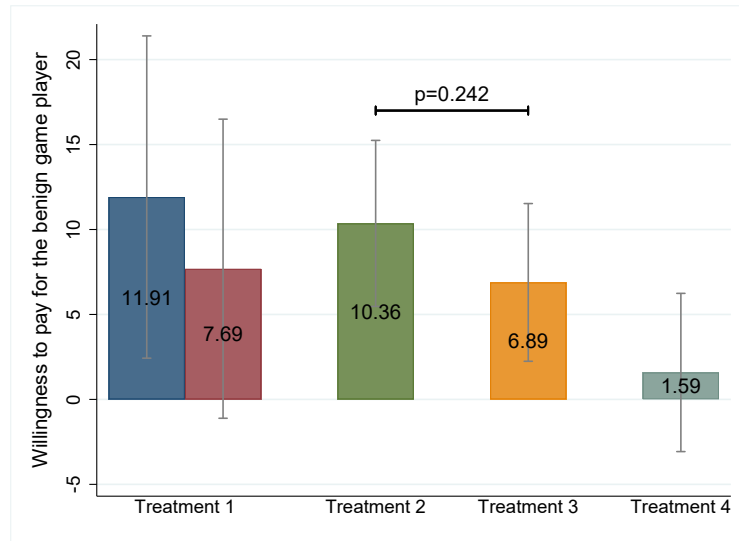
Notes: Figure plots the average benign premiums in treatments 2 to 4, depending on their malign-game partners' actions.

Figure 3.3: Benign Premiums depending on Malign-game Partner's Action

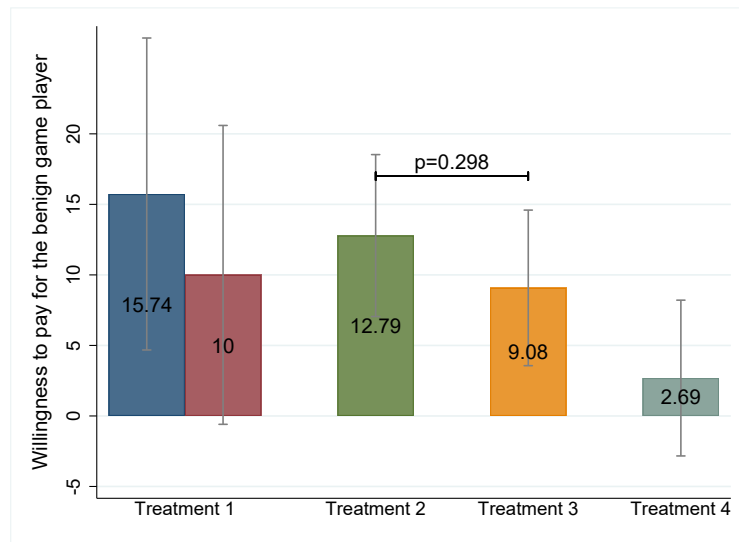


Note: Figure plots the average money subjects shared in the dictator game at stage one depending on their own actions in the malign game.

Figure 3.4: The Correlation between Action in the Malign Game and How Much shared in DG



Panel (a)



Panel (b)

Note: Figure plots the willingness to pay towards the benign-game player using alternative coding methods. Panel (a) measures the benign premium as the minimum value of the interval from the multiple price list. Panel (b) measures the benign premium as the maximum value of the interval from the multiple price list.

Figure 3.5: Robustness Check - Benign Premium across Treatments

3.7 Tables

Table 3.1: The Benign and Malign Games

Harmony Game			Prisoners' Dilemma		
	C	D		C	D
C	40,40	10,30	C	40,40	20,120
D	30,10	0,0	D	120,20	30,30

Note: The harmony game is the benign game, and the prisoner's dilemma is the malign game.

Table 3.2: The Multiple Price List

	Option 1	Option 2
Choice 1	Amount transferred to me by B	Amount transferred to me by D+10
Choice 2	Amount transferred to me by B	Amount transferred to me by D+20
Choice 3	Amount transferred to me by B	Amount transferred to me by D+30
Choice 4	Amount transferred to me by B	Amount transferred to me by D+40
Choice 5	Amount transferred to me by B	Amount transferred to me by D+50
Choice 6	Amount transferred to me by B	Amount transferred to me by D+60
Choice 7	Amount transferred to me by B	Amount transferred to me by D+70
Choice 8	Amount transferred to me by B	Amount transferred to me by D+80
Choice 9	Amount transferred to me by B	Amount transferred to me by D+90
Choice 10	Amount transferred to me by B	Amount transferred to me by D+100

Note: Table shows the multiple price list shown to subject A if she chooses B over D in the first choice.

Table 3.3: Benign Premium across Treatments

Treatment		Obs	WTP	P-value	P-value
				$H_0 : wtp = 0$	$H_0 : wtp_{Tx} = wtp_{T3}$
Treatment 1	benignP VS stranger	68	13.82	0.009	
	stranger VS malignP	65	8.85	0.072	
Treatment 2		251	11.57	0.000	0.298
Treatment 3		228	7.98	0.002	
Treatment 4		227	2.14	0.401	0.077

Notes: The first row in treatment 1 represents the WTP for the benign game player versus a stranger, and the second row in treatment 1 represents the WTP for a stranger versus a malign game player. Column (3) reports the p-value of t-tests. Column (4) reports the p-value of Ranksum tests.

Table 3.4: Summary Statistics

Variable	Treatment 1	Treatment 2	Treatment 3	Treatment 4
How much kept in DG	130.7 (40.66)	131.9 (41.33)	133.6 (42.24)	134.7 (42.64)
Cooperation rate in the malign game	0.400 (0.494)	0.413 (0.494)	0.360 (0.481)	0.401 (0.491)
Share of \geq college degree	0.778 (0.418)	0.813 (0.391)	0.843 (0.365)	0.783 (0.413)
Income	3.778 (1.538)	3.924 (1.642)	3.863 (1.541)	3.826 (1.657)
Share of female	0.547 (0.500)	0.587 (0.493)	0.623 (0.486)	0.527 (0.501)
Age	37.23 (9.955)	38.54 (11.42)	37.28 (10.02)	38.81 (11.79)
Share of working people	0.794 (0.406)	0.809 (0.394)	0.814 (0.391)	0.792 (0.407)
Observations	133	251	228	227

Note: Standard deviations in parentheses.

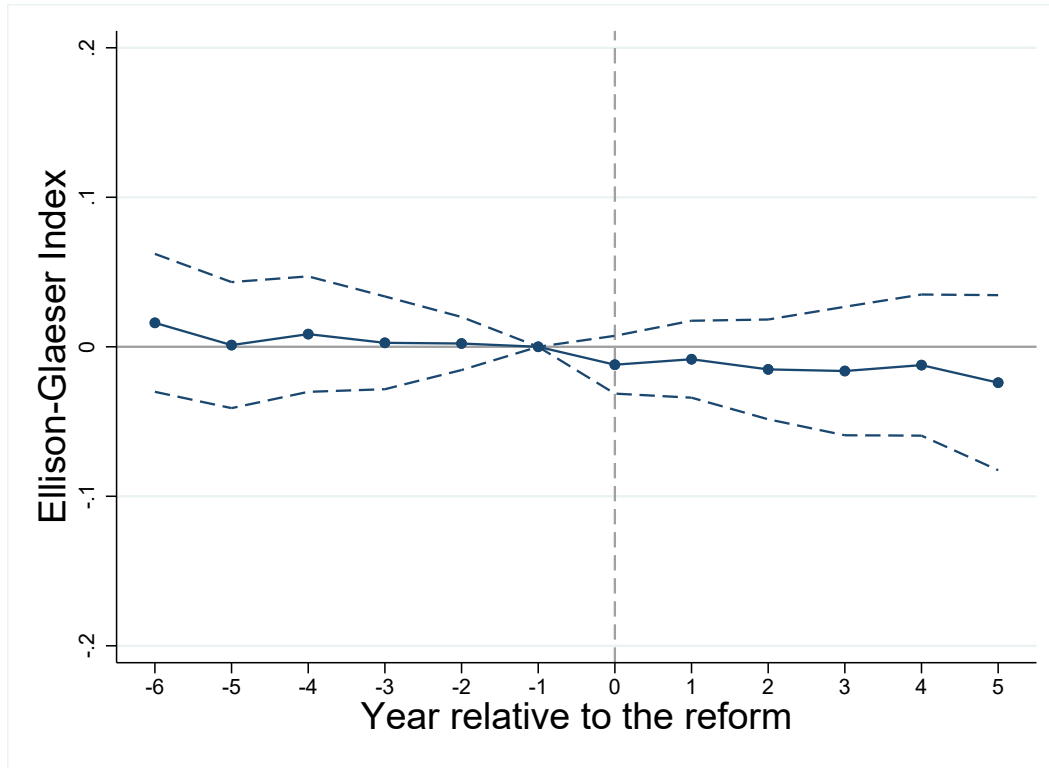
Table 3.5: Benign Premiums in Treatment 2

	Malign-game player		Ranksum Test
	Cooperate	Defect	p-value
Benign game only ^a	13.89	26.84	0.065
Malign game only ^b	-14.25	12.79	0.0001
Mean	-0.047	20.21	0.0001

Notes: a - subjects who only played the benign game and observe the
malign game. b - subjects who only played the malign game.

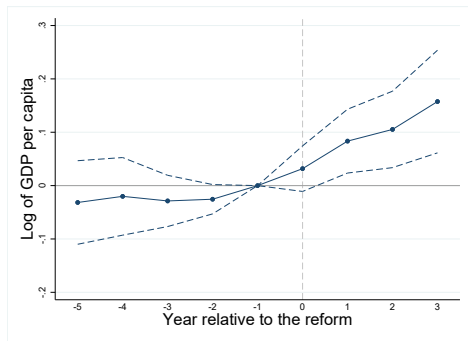
4.0 Appendices

4.1 Appendix for Chapter 1

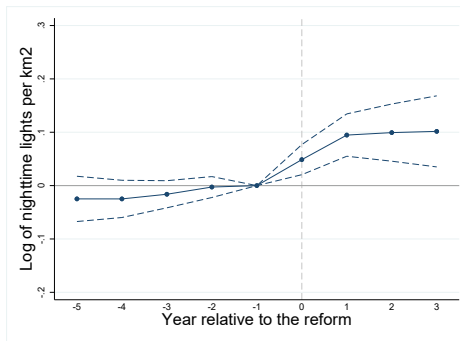


Notes: Figure plots estimates of the effect of the incorporation reform on geographical concentration index in treated prefectures in the years before and after the reform, based on estimates of coefficients from equation 1.2 at the prefecture level. The dependent variable is the Ellison-Glaeser index. For prefectures that had several incorporations, I only consider the first incorporation. Dashed line is 95 percent confidence interval for the outcome (solid series).

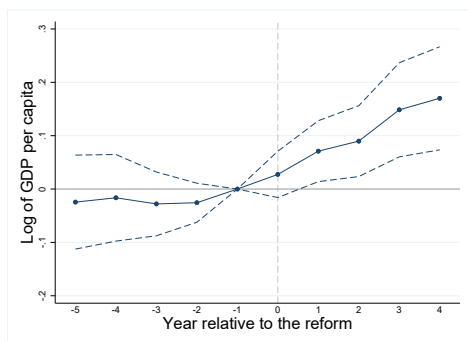
Figure 4.1: Event Study of the Reform on Geographical Concentration



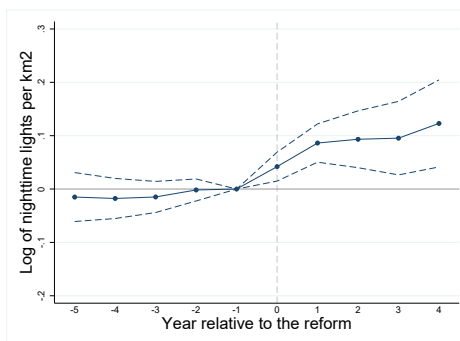
(a) The impact on GDP per capita (3-year gap)



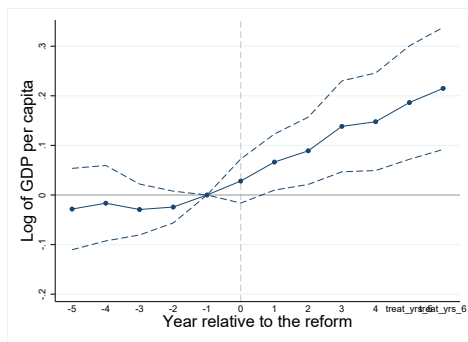
(b) The Impact on Lights per km² (3-year gap)



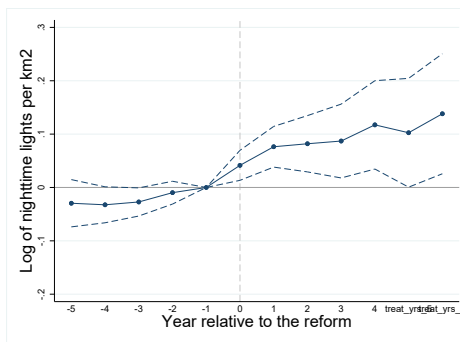
(c) The impact on GDP per capita (4-year gap)



(d) The Impact on Lights per km² (4-year gap)



(e) The impact on GDP per capita (6-year gap)



(f) The Impact on Lights per km² (6-year gap)

Notes: Robustness check for Approach II. I compare counties that experience the current incorporation to counties that would experience the reform three, four, six and seven years later respectively. Figure plots estimates of the effect of the reform on GDP and nighttime lights in treated counties in the years before and after the reform, based on estimates of coefficients from equation 1.3. The dependent variables are the log of GDP per capita or the log of nighttime lights per km². Dashed line is 95 percent confidence interval for outcomes (solid series).

Figure 4.2: Robustness: The Impact of Market Integration (Approach II)

Table 4.1: Factors that Predict Timing of Incorporations

	Timing of incorporations			
	1998	2002	2006	2011
	(1)	(2)	(3)	(4)
Population (lag)	0.000 (0.000)	-0.001 (0.002)	0.003 (0.002)	0.002 (0.008)
Manufacturing share of GDP (lag)	-0.217 (0.222)	0.784 (1.284)	0.026 (1.051)	7.005 (7.931)
Tertiary share of GDP (lag)	-0.246 (0.259)	0.681 (1.361)	1.154 (1.043)	11.103 (9.810)
Ratio of gov. expenditure to gov. revenue (lag)	-0.002 (0.011)	-0.058 (0.101)	-0.132 (0.143)	-0.264 (0.905)
Ratio of gov. revenue to GDP (lag)	0.559 (1.213)	1.463 (10.764)	-31.455 (18.889)	1.660 (54.023)
Ratio of gov. expenditure to GDP (lag)	-0.462 (0.873)	2.722 (7.432)	16.192* (8.140)	18.835 (43.958)
Log of lights per km ² (lag)	0.042 (0.042)	0.203* (0.119)	0.323* (0.158)	-0.581 (0.626)
Dummy of provincial capital	-0.003 (0.012)	-0.104 (0.256)	0.651 (0.417)	-0.257 (0.874)
Dummy of direct-administered municipalities of China	0.101 (0.101)	0.033 (0.175)	1.304*** (0.219)	- -
Observations	63	41	21	13

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses.

Table 4.2: Estimated Effects of the Reform on Economic Growth:
Approach I (Without Sample of Direct-administered Municipalities of
China)

Dependent variable	Log of GDP per capita		Log of lights per km^2	
	(1)	(2)	(3)	(4)
Reform	0.160*** (0.041)	0.147*** (0.038)	0.065** (0.027)	0.044* (0.026)
County-level controls		Y		Y
County FE	Y	Y	Y	Y
Province×Year FE	Y	Y	Y	Y
Observations	4,169	3,988	4,370	3,982
R-squared	0.964	0.970	0.983	0.984
Mean DV	8.968	8.968	1.696	1.724
Std.Dev. DV	0.913	0.913	0.853	0.831

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The columns presents estimates of β_1 from equation 1.1. All regressions include a full set of county and province×year fixed effect. Robust standard errors are in parentheses, clustered at the county level. The county-level controls include manufacturing share of GDP, tertiary industry share of GDP, ratio of government expenditure to government revenue. Log of population is also included as the county-level control for the results on nighttime lights.

Table 4.3: Estimated Effects of the Reform on Economic Growth:
Approach II (Without Sample of Direct-administered Municipalities of
China)

Dependent variable	Log of GDP per capita		Log of lights per km ²	
	(1)	(2)	(3)	(4)
Treatment×Post	0.125** (0.047)	0.136*** (0.050)	0.109*** (0.029)	0.058** (0.027)
County-level controls		Y		Y
County FE	Y	Y	Y	Y
Province×Year FE	Y	Y	Y	Y
Observations	9,618	9,367	9,635	9,367
R-squared	0.977	0.983	0.987	0.989
Mean DV	8.761	8.790	1.784	1.812
Std.Dev. DV	0.654	0.634	0.679	0.658

Note: *** p<0.01, ** p<0.05, * p<0.1. The columns presents estimates of β_1 from equation 1.4. All regressions include a full set of county and province×year fixed effect. Robust standard errors are in parentheses, clustered at the incorporation level. The county-level controls include manufacturing share of GDP, tertiary industry share of GDP, ratio of government expenditure to government revenue. Log of population is also included as the county-level control for the results on night-time lights.

4.2 Appendix for Chapter 2

4.2.1 Fertility Penalty Data

The fertility penalty data are taken from Scharping (2013), which provides an overall view of China's fertility policies and outcomes. Scharping draws on a large number of primary and secondary sources (statistics, laws, directives, internal documents, conferences, etc.) at local, national and international levels, collected over 10 years. Specifically, he documented the complete record of the published fine rates across provinces ranging from 1979 to 2000.¹ We modified the penalty data for seven provinces: Beijing, Inner Mongolia, Liaoning, Heilongjiang, Jiangxi, Shandong and Guangxi. Furthermore, we extend the fine rates data from 2000 to 2010 using provincial governments' documents.²

There were mainly three forms of fines as documented by Scharping (2013). We transform all three forms into a percentages of a household's annual income. The first type was collected from wage earners in the form of regular deductions. For fines levied as wage deductions, we follow Ebenstein (2010)'s method by calculating the present value of the penalty at a 2 percent discount rate. For example, in February 1980 Guangdong province ratified a fine of 10 percent of wage from each parent for 14 years for an unsanctioned birth. The present value of the fine is then 1.21 years of income. The detailed calculation is as follows

$$Penalty_{Guangdong}^{1980} = 0.1 + 0.1 \times (1 - 0.02) + 0.1 \times (1 - 0.02)^2 + \dots + 0.1 \times (1 - 0.02)^{13} = 1.21 \quad (4.2.1)$$

The second type of fines was levied as a share of annual income that needs to be paid in a single payment. For example, Shanghai employed a rule in 1992 that an unauthorized birth carried an immediate payment of three years of household income. In this case, no transformation is needed.

¹The data source of Scharping (2013) is based on two books: Zhongguo Jihua Shengyu Quanshu [Encyclopedia of Birth Planning in China] and Zhongguo Renkou Congshu 1987-1993.

²The data were downloaded from <http://www.pkulaw.cn/>

The third type of fines was collected as an immediate payment of a certain amount of money, independent of one's household income. For example, from 1995 to 2000, Guangxi ratified the fine as an amount between 2,000 and 50,000 yuan. In this circumstance, Ebenstein (2010) calculates the fine amount with the following assumptions. First, the fine was collected at the maximum amount of the range;³ second, the average household annual income was fixed at 10,000 RMB across province and time.⁴ We can apply his rules to the Guangxi example. According to the first assumption, the maximum amount in the range, 50,000 yuan, is taken as the penalty amount. Then based on the second assumption, the 50,000 yuan fine is equivalent to $\frac{50,000}{10,000} = 5$ years of household income.

For the third type of fines, instead of assuming that the annual household income was fixed at 10,000 RMB across province and year, we impute the fine into a share of household income using the provincial average household annual income in a certain year. The income is only averaged at the provincial level because fines data are only available at this level, and we want to keep the units of data consistent. The income data are taken from the China Statistical Yearbooks. We add this variation of income across province and time for two reasons. First, there is an substantial variation of annual household income across provinces in China at that time. For instance, the annual household incomes were 4,630 yuan in Liaoning and 7,095 yuan in Beijing in 1993. In the same year, both provinces levied an amount of 50,000 yuan fine for an unauthorized child. Clearly, the same amount of fine would not be the same for households in Liaoning and households in Beijing.

Second, there is also substantial time variation in annual household income within a province due to the rapid economic growth in the 1990s. For example, Beijing ratified the fine as 50,000 yuan from 1991 to 2000, while the average household annual income increased from 4,371 yuan in 1991 to 20,833 yuan in 2000. Hence, we calculate the fines as $\frac{50,000}{4,371} = 11.44$ years of income in 1991, and as $\frac{50,000}{20,833} = 2.40$ years of income in 2000.

³We also assume that the fine was collected at the maximum amount, hereafter we only talk about the maximum amount.

⁴The only exception for the second assumption is Heilongjiang from 1983 to 1988, whose average annual household income is taken as 1,000 yuan.

4.2.2 The One-tailed T Test for Sex Ratios

The publicly available data of sex ratios at the first birth in urban China is very limited: we have one year across provinces' sex ratios from the 2000 Census. The small sample size leads to a lot of random variations in the sex ratios. We cannot simply say provinces with sex ratios that are out of the normal range have severe sex selections.

To address this concern, we construct a one-tailed t-statistics to test whether the calculated sex ratio (with limited sample size) is statistically different from the biological sex ratio (1.06 boys/girls). We treat the gender of the first child, D_i as a Bernoulli trial with a probability $p = 0.515(106/206)$ to be a boy. Let's denote child i 's gender as D_i . Formally,

$$D_i = \begin{cases} 1, & \text{if } i \text{ is a boy.} \\ 0, & \text{otherwise.} \end{cases} \quad (4.2.2)$$

Under biological sex ratio, we should observe a boy with a probability of 0.515. Let the biological share of boys be $\mathcal{P}(D_i = 1) = p$, the calculated share of boys be \hat{p} , and the sample size be n . The null hypothesis is $H_0 : p = 0.515$. As D_i is a Bernoulli trial,

$$\hat{p} = \frac{\sum_{i=1}^n D_i}{n}. \quad (4.2.3)$$

Then, we can get $E(D_i) = p$ and $Var(D_i) = p(1 - p)$. By the central limit theorem,

$$\sqrt{n}(\hat{p} - p) \equiv \frac{\sqrt{n} \sum_{i=1}^n D_i}{n} - p \xrightarrow{d} \mathbf{N}(0, p(1 - p)) \quad (4.2.4)$$

As a result,

$$\frac{\sqrt{n}(\hat{p} - p)}{\sqrt{p(1 - p)}} \xrightarrow{d} \mathbf{N}(0, 1) \quad (4.2.5)$$

So we can construct the following t-statistics:

$$t \equiv \frac{\sqrt{n}(\hat{p} - 0.515)}{\sqrt{\hat{p}(1 - \hat{p})}}. \quad (4.2.6)$$

Since we are only concerned with the possibility that people endogenously choose boys over girls, we focus on a one-tailed t test and reject the null hypothesis if $t > 1.645$. We find that six provinces' sex ratios at first birth were significantly higher than the normal sex ratio at the 5 percent level: Beijing, Jiangsu, Jiangxi, Hubei, Guangdong, and Guangxi (Table 4.4).

Table 4.4: Sex Ratio at First Birth (Males/Females) in
Urban China in 2000

Province	Obs	Sex ratio	One-tailed t-statistics
Beijing	4579	1.130	2.156
Tianjin	2282	1.063	0.074
Hebei	8356	1.090	1.277
Shanxi	5582	1.081	0.742
Liaoning	10525	1.061	0.063
Jilin	4940	1.104	1.426
Heilongjiang	7108	1.079	0.747
Shanghai	5518	1.089	1.014
Jiangsu	12635	1.099	2.039
Zhejiang	8938	1.072	0.526
Anhui	6499	1.082	0.816
Fujian	5182	1.091	1.044
Jiangxi	4367	1.178	3.489
Shandong	17607	1.080	1.253
Henan	11090	1.076	0.808
Hubei	9403	1.124	2.821
Hunan	7072	1.052	-0.309
Guangdong	20607	1.166	6.825
Guangxi	4418	1.131	2.160
Chongqing	4010	1.022	-1.150
Sichuan	8763	1.098	1.644
Guizhou	3886	0.988	-2.201
Yunnan	4056	1.004	-1.729
Shaanxi	4365	1.060	-0.002
Gansu	3240	1.058	-0.042

Data source: China Census 10% sample (2000)

Table 4.5: Estimates of OCP Exposure on Trust in Phase II
(1991-2010): Excluding High Sex Ratio Provinces

Dependent variable:	Trust in			
	Neighbors		Local governments	
	(1)	(2)	(3)	(4)
OCP exposure	0.038 (0.063)	0.021 (0.062)	-0.039 (0.097)	-0.059 (0.101)
Firstborn daughter ×OCP exposure	-0.306*** (0.087)	-0.304*** (0.085)	-0.111 (0.172)	-0.108 (0.180)
p-Value	[0.003]	[0.002]	[0.527]	[0.557]
Bootstrap p-Value	[0.004]	[0.004]	[0.692]	[0.784]
Individual controls		Y		Y
Cohort FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
Mean DV	6.409	6.409	4.218	4.218
Std.Dev.DV	2.059	2.059	2.500	2.500
Observations	1,472	1,472	1,472	1,472
R-squared	0.105	0.119	0.091	0.103

*** p<0.01, ** p<0.05, * p<0.1. OCP exposure in phase II is defined as the five-year mean value of the fertility penalty rates in province p after individual i had his/her first child. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. We use a wild cluster bootstrap-t procedure that are clustered at the province level for improved inference with a small number of clusters (Cameron, Gelbach and Miller, 2008). We report the corresponding p-values in brackets. We also report the p-values for OLS with clustered data. Number of clusters: 25.

Table 4.6: Estimates of OCP Exposure on Trust in Phase II

Dependent variable:	Trust in			
	Neighbors		Local governments	
	(1)	(2)	(3)	(4)
OCP exposure (min)	0.077 (0.078)	0.068 (0.074)	-0.048 (0.063)	-0.058 (0.066)
Firstborn daughter × OCP exposure (min)	-0.220*** (0.073)	-0.216*** (0.071)	-0.017 (0.122)	-0.010 (0.126)
Mean DV	6.407	6.407	4.304	4.304
Std.Dev.DV	2.065	2.065	2.470	2.470
Observations	1,897	1,897	1,897	1,897

*** p<0.01, ** p<0.05, * p<0.1. OCP exposure is defined as the minimum of the five-year fertility penalty rates in province p after individual i had his/her first child. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. Number of clusters: 25.

Table 4.7: Estimates of OCP Exposure on Trust in Phase I

Dependent variable:	Trust in			
	Neighbors		Local governments	
	(1)	(2)	(3)	(4)
OCP exposure (min)	0.078 (0.068)	0.072 (0.069)	0.165** (0.059)	0.151** (0.064)
Firstborn daughter ×OCP exposure (min)	-0.022 (0.100)	0.006 (0.103)	-0.227*** (0.053)	-0.212*** (0.062)
Mean DV	6.582	6.582	4.957	4.957
Std.Dev.DV	2.072	2.072	2.502	2.502
Observations	822	822	822	822

*** p<0.01, ** p<0.05, * p<0.1. OCP exposure is defined as the minimum of the five-year family-planning rate in province p after an individual had his/her first child. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. Number of clusters: 25.

Table 4.8: Estimates of OCP Exposure on Trust in Phase II

Dependent variable:	Trust in			
	Neighbors		Local governments	
	(1)	(2)	(3)	(4)
OCP exposure (4-year mean)	0.072 (0.062)	0.060 (0.059)	-0.009 (0.080)	-0.022 (0.087)
Firstborn daughter × OCP exposure (4-year mean)	-0.303*** (0.087)	-0.296*** (0.087)	-0.102 (0.162)	-0.100 (0.170)
Individual controls		Y		Y
Cohort FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
Mean DV	6.407	6.407	4.304	4.304
Std.Dev.DV	2.065	2.065	2.470	2.470
Observations	1,897	1,897	1,897	1,897
R-squared	0.089	0.098	0.078	0.088

*** p<0.01, ** p<0.05, * p<0.1. OCP exposure is defined as the four-year mean value of the fertility penalty rates in province p after individual i had his/her first child. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. Number of clusters: 25.

Table 4.9: Estimates of OCP Exposure on Trust in Phase II

Dependent variable:	Trust in			
	Neighbors		Local governments	
	(1)	(2)	(3)	(4)
OCP exposure (6-year mean)	0.036 (0.069)	0.023 (0.066)	-0.076 (0.074)	-0.088 (0.076)
Firstborn daughter × OCP exposure (6-year mean)	-0.273*** (0.082)	-0.266*** (0.083)	-0.032 (0.136)	-0.033 (0.141)
Individual controls		Y		Y
Cohort FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
Mean DV	6.407	6.407	4.304	4.304
Std.Dev.DV	2.065	2.065	2.470	2.470
Observations	1,897	1,897	1,897	1,897
R-squared	0.088	0.097	0.078	0.088

*** p<0.01, ** p<0.05, * p<0.1. OCP exposure is defined as the six-year mean value of the fertility penalty rates in province p after individual i had his/her first child. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. Number of clusters: 25.

Table 4.10: Estimates of OCP Exposure on Trust in Phase I

Dependent variable:	Trust in			
	Neighbors		Local governments	
	(1)	(2)	(3)	(4)
OCP exposure (4-year mean)	0.075 (0.061)	0.064 (0.060)	0.088 (0.066)	0.082 (0.059)
Firstborn daughter ×OCP exposure (4-year mean)	-0.065 (0.072)	-0.043 (0.071)	-0.250*** (0.060)	-0.248*** (0.067)
Individual controls		Y		Y
Cohort FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
Mean DV	6.582	6.582	4.957	4.957
Std.Dev.DV	2.072	2.072	2.502	2.502
Observations	822	822	822	822
R-squared	0.141	0.155	0.139	0.155

*** p<0.01, ** p<0.05, * p<0.1. OCP exposure is defined as the four-year average family-planning rate in province p after an urban resident had his/her first child. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. Number of clusters: 25.

Table 4.11: Estimates of OCP Exposure on Trust in Phase I

Dependent variable:	Trust in			
	Neighbors		Local governments	
	(1)	(2)	(3)	(4)
OCP exposure (6-year mean)	0.144 (0.111)	0.127 (0.108)	0.152 (0.117)	0.138 (0.104)
Firstborn daughter ×OCP exposure (6-year mean)	-0.075 (0.107)	-0.037 (0.104)	-0.439*** (0.102)	-0.441*** (0.110)
Individual controls		Y		Y
Cohort FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
Mean DV	6.582	6.582	4.957	4.957
Std.Dev.DV	2.072	2.072	2.502	2.502
Observations	822	822	822	822
R-squared	0.143	0.157	0.140	0.157

*** p<0.01, ** p<0.05, * p<0.1. OCP exposure is defined as the six-year average family-planning rate in province p after an urban resident had his/her first child. All regressions include a full set of province and cohort fixed effects. In parentheses are standard errors clustered by province. We also report the p-values for OLS with clustered data. Number of clusters: 25.

Table 4.12: Distribution of Responses: Trust in Local Governments

	Trust in local governments	
	First child was a boy	First child was a girl
0	9.87	9.73
1	6.65	5.42
2	8.86	9.85
3	10.57	12.94
4	5.94	5.53
5	29.81	31.97
6	9.77	8.41
7	7.25	5.75
8	7.05	6.42
9	2.01	2.32
10	2.22	1.66
Observations.	993	904
Mean DV	4.35	4.25
Std.Dev. DV	2.33	2.30

The table shows the distribution of responses to the question regarding trust in local government officials, split by the gender of the first child. Data source: CFPS (2016).

Table 4.13: Correlation between Fertility
Penalty and Family-planning Rate

Dependent variable: Family-planning rate	
Fertility penalty	1.588** (0.659)
Province FE	Y
Year FE	Y
Observations	142

*** p<0.01, ** p<0.05, * p<0.1 Standard errors in parentheses. The sample is a province-level panel (unbalanced) from 1979 to 1987, when both fertility penalty and family-planning rate are available.

5.0 References

- Akerberg, Daniel, Kevin Caves, and Garth Frazer**, “Structural identification of production functions,” *Working Paper*, 2006.
- Ambuehl, Sandro and Shengwu Li**, “Belief updating and the demand for information,” *Games and Economic Behavior*, 2018, 109, 21–39.
- Andreoni, James, Nikos Nikiforakis, and Jan Stoop**, “Are the rich more selfish than the poor, or do they just have more money? A natural field experiment,” Technical Report, National Bureau of Economic Research 2017.
- Bai, Chong-En, Yingjuan Du, Zhigang Tao, and Sarah Y Tong**, “Local protectionism and regional specialization: evidence from China’s industries,” *Journal of International Economics*, 2004, 63 (2), 397–417.
- Banister, Judith**, *China’s changing population*, Stanford University Press, 1991.
- Barreda-Tarrazona, Iván, Ainhua Jaramillo-Gutiérrez, Marina Pavan, and Gerardo Sabater-Grande**, “Individual characteristics vs. experience: an experimental study on cooperation in prisoner’s dilemma,” *Frontiers in psychology*, 2017, 8, 596.
- Bartling, BJÖRN and URS Fischbacher**, “Shifting the blame: on delegation and responsibility,” *The Review of Economic Studies*, 2012, 79 (1), 67–87.
- Barwick, Panle Jia, Shengmao Cao, and Shanjun Li**, “Local protectionism, market structure, and social welfare: China’s automobile market,” *NBER Working Paper No. 23678*, 2017.
- Benjamin, Daniel J**, “Errors in probabilistic reasoning and judgment biases,” in “Handbook of Behavioral Economics: Applications and Foundations 1,” Vol. 2, Elsevier, 2019, pp. 69–186.

- Bernstein, Thomas P and Xiaobo Lü**, “Taxation without representation: peasants, the central and the local states in reform China,” *The China Quarterly*, 2000, 163, 742–763.
- Bó, Ernesto Dal, Pedro Dal Bó, and Erik Eyster**, “The demand for bad policy when voters underappreciate equilibrium effects,” *The Review of Economic Studies*, 2018, 85 (2), 964–998.
- Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller**, “Bootstrap-based improvements for inference with clustered errors,” *The Review of Economics and Statistics*, 2008, 90 (3), 414–427.
- Cantoni, Davide, Yuyu Chen, David Y Yang, Noam Yuchtman, and Y Jane Zhang**, “Curriculum and ideology,” *Journal of Political Economy*, 2017, 125 (2), 338–392.
- Cartier, Carolyn**, “A political economy of rank: The territorial administrative hierarchy and leadership mobility in urban China,” *Journal of Contemporary China*, 2016, 25 (100), 529–546.
- Cason, Timothy N, Sau-Him Paul Lau, and Vai-Lam Mui**, “Prior interaction, identity, and cooperation in the Inter-group Prisoner’s Dilemma,” *Journal of Economic Behavior & Organization*, 2019.
- Chen, Yuyu and David Y Yang**, “Historical traumas and the roots of political distrust: Political inference from the great Chinese famine,” *Working Paper*, 2019.
- , **Hongbin Li, and Lingsheng Meng**, “Prenatal sex selection and missing girls in China: Evidence from the diffusion of diagnostic ultrasound,” *Journal of Human Resources*, 2013, 48 (1), 36–70.
- Christian, Charles W**, “Voluntary compliance with the individual income tax: results from the 1988 TCMP study,” *The IRS Research Bulletin*, 1994, 1500 (9-94), 35–42.
- Coffman, Lucas C**, “Intermediation reduces punishment and reward,” *American Economic Journal: Microeconomics*, 2011, 3 (4), 77–106.

- Cooper, Russell, Douglas V DeJong, Robert Forsythe, and Thomas W Ross**, “Cooperation without reputation: Experimental evidence from prisoner’s dilemma games,” *Games and Economic Behavior*, 1996, 12 (2), 187–218.
- Cox, Dennis**, “Raising revenue in the underground economy,” *National Tax Journal*, 1984, 37 (3), 283–288.
- Davoodi, Hamid and Heng fu Zou**, “Fiscal decentralization and economic growth: A cross-country study,” *Journal of Urban economics*, 1998, 43 (2), 244–257.
- Deshpande, Manasi and Yue Li**, “Who is screened out? application costs and the targeting of disability programs,” *American Economic Journal: Economic Policy*, 2019, 11 (4), 213–48.
- Donaldson, Dave**, “The gains from market integration,” *The Annual Review of Economics*, 2015, 7 (1), 619–647.
- , “Railroads of the Raj: Estimating the impact of transportation infrastructure,” *American Economic Review*, 2018, 108 (4-5), 899–934.
- **and Richard Hornbeck**, “Railroads and American economic growth: A “market access” approach,” *The Quarterly Journal of Economics*, 2016, 131 (2), 799–858.
- Ebenstein, Avraham**, “The “missing girls” of China and the unintended consequences of the one child policy,” *Journal of Human Resources*, 2010, 45 (1), 87–115.
- Edlund, Lena**, “Son preference, sex ratios, and marriage patterns,” *Journal of Political Economy*, 1999, 107 (6), 1275–1304.
- Edwards, Ward**, “Conservatism in human information processing,” in Kleinmuntz B, ed., *Formal Representation of Human Judgement*, New York: Wiley, 1968, p. 17–52.
- Ellison, Glenn and Edward L Glaeser**, “Geographic concentration in US manufacturing industries: a dartboard approach,” *Journal of Political Economy*, 1997, 105 (5), 889–927.

- Eyer, Jonathan and Matthew E Kahn**, “Prolonging coal’s sunset: The causes and consequences of local protectionism for a declining polluting industry,” *NBER Working Paper No. 23190*, 2017.
- Eyster, Erik and Matthew Rabin**, “Cursed equilibrium,” *Econometrica*, 2005, 73 (5), 1623–1672.
- Faber, Benjamin**, “Trade integration, market size, and industrialization: evidence from China’s National Trunk Highway System,” *Review of Economic Studies*, 2014, 81 (3), 1046–1070.
- Fadlon, Itzik and Torben Heien Nielsen**, “Family Labor Supply Responses to Severe Health Shocks,” *NBER Working Paper No. 21352*, 2015.
- Fiorina, Morris P**, “Legislator uncertainty, legislative control, and the delegation of legislative power,” *Journal of Law, Economics, and Organization*, 1986, 2, 33.
- Fitzpatrick, Sheila**, *Everyday stalinism: Ordinary life in extraordinary times: Soviet Russia in the 1930s*, Oxford University Press, 1999.
- Fong, Mei**, *One child*, Oneworld Publications, 2016.
- Gemmell, Norman, Richard Kneller, and Ismael Sanz**, “Fiscal decentralization and economic growth: spending versus revenue decentralization,” *Economic Inquiry*, 2013, 51 (4), 1915–1931.
- Gilbert, Daniel T and Patrick S Malone**, “The correspondence bias,” *Psychological bulletin*, 1995, 117 (1), 21.
- Gupta, Monica Das, Zhenghua Jiang, Bohua Li, Zhenming Xie, Chung Woojin, and Hwa-Ok Bae**, “Why is son preference so persistent in east and south Asia? A cross-country study of China, India and the Republic of Korea,” *Journal of Development Studies*, December 2003, 40 (2), 153–187.

- Guryan, Jonathan**, “Desegregation and black dropout rates,” *American Economic Review*, 2004, 94 (4), 919–943.
- Haggag, Kareem, Devin G Pope, Kinsey B Bryant-Lees, and Maarten W Bos**, “Attribution bias in consumer choice,” *The Review of Economic Studies*, 2019, 86 (5), 2136–2183.
- Hamman, John R, George Loewenstein, and Roberto A Weber**, “Self-interest through delegation: An additional rationale for the principal-agent relationship,” *The American Economic Review*, 2010, 100 (4), 1826–1846.
- Han, Li and James Kai-Sing Kung**, “Fiscal incentives and policy choices of local governments: Evidence from China,” *Journal of Development Economics*, 2015, 116, 89–104.
- Henderson, J Vernon, Adam Storeygard, and David N Weil**, “Measuring economic growth from outer space,” *American Economic Review*, 2012, 102 (2), 994–1028.
- Hodler, Roland and Paul A Raschky**, “Regional favoritism,” *The Quarterly Journal of Economics*, 2014, 129 (2), 995–1033.
- Holz, Carsten A**, “No razor’s edge: Reexamining Alwyn Young’s evidence for increasing interprovincial trade barriers in China,” *The Review of Economics and Statistics*, 2009, 91 (3), 599–616.
- Huang, Wei, Xiaoyan Lei, and Ang Sun**, “The great expectations: Impact of one-child policy on education of girls,” *IZA Discussion Paper*, 2015.
- Jin, Hehui, Yingyi Qian, and Barry R Weingast**, “Regional decentralization and fiscal incentives: Federalism, Chinese style,” *Journal of Public Economics*, 2005, 89 (9-10), 1719–1742.
- Johnson, Eric A**, *Nazi terror: The gestapo, Jews, and ordinary Germans*, Basic Books, 2000.
- Johnson, Noel D. and Alexandra Mislin**, “How much should we trust the world values survey trust question?,” *Economics Letters*, 2012, 116 (2), 210–212.

- Jones, Edward E and Victor A Harris**, “The attribution of attitudes,” *Journal of experimental social psychology*, 1967, 3 (1), 1–24.
- Kahneman, Daniel and Amos Tversky**, “Subjective probability: A judgment of representativeness,” *Cognitive psychology*, 1972, 3 (3), 430–454.
- **and** – , “On the psychology of prediction.,” *Psychological review*, 1973, 80 (4), 237.
- King, Gary, Jennifer Pan, and Margaret E Roberts**, “How censorship in China allows government criticism but silences collective expression,” *American Political Science Review*, 2013, 107 (2), 326–343.
- leung Chan, Alan Kam**, *Mencius: contexts and interpretations*, University of Hawaii Press, 2002.
- Li, Hongbin and Junsen Zhang**, “Do high birth rates hamper economic growth?,” *The Review of Economics and Statistics*, 2007, 89 (1), 110–117.
- **and Li-An Zhou**, “Political turnover and economic performance: the incentive role of personnel control in China,” *Journal of Public Economics*, 2005, 89 (9-10), 1743–1762.
- , **Junjian Yi, and Junsen Zhang**, “Estimating the effect of the one-child policy on the sex ratio imbalance in China: Identification based on the difference-in-differences,” *Demography*, 2011, 48 (4), 1535–1557.
- Li, Lixing and Xiaoyu Wu**, “Gender of children, bargaining power, and intrahousehold resource allocation in China,” *Journal of Human Resources*, 2011, 46 (2), 295–316.
- Lichter, Andreas, Max Loeffler, and Sebastian Siegloch**, “The Economic costs of mass surveillance: Insights from Stasi spying in East Germany,” *IZA Discussion Paper*, 2015.
- Liu, Xiuyan, Jiangnan Zeng, and Qiyao Zhou**, “The chosen fortunate in the urbanization process in China? Evidence from a geographic regression discontinuity study,” *Review of Development Economics*, 2019.

- Long, Cheryl Xiaoning and Jun Wang**, “Judicial local protectionism in China: An empirical study of IP cases,” *International Review of Law and Economics*, 2015, 42, 48–59.
- Lowe, Matt**, “Types of contact: A field experiment on collaborative and adversarial caste integration,” 2020.
- Lu, Warren Wenzhi and Kellee S Tsai**, “Inter-Governmental Vertical Competition in China’s Urbanization Process,” *Journal of Contemporary China*, 2019, 28 (115), 99–117.
- Mathews, T. J., Brady E. Hamilton, and others**, “Trend analysis of the sex ratio at birth in the United States,” *National Vital Statistics Reports*, 2005, 53 (20), 1–17.
- Melitz, Marc J**, “The impact of trade on intra-industry reallocations and aggregate industry productivity,” *Econometrica*, 2003, 71 (6), 1695–1725.
- Möbius, Markus M, Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat**, “Managing self-confidence,” *NBER Working Paper No. 17014*, 2014.
- Nunn, Nathan and Leonard Wantchekon**, “The slave trade and the origins of mistrust in Africa,” *The American Economic Review*, 2011, 101 (7), 3221–3252.
- Oexl, Regine and Zachary J Grossman**, “Shifting the blame to a powerless intermediary,” *Experimental Economics*, 2013, 16 (3), 306–312.
- Peysakhovich, Alexander and David G Rand**, “Habits of virtue: Creating norms of cooperation and defection in the laboratory,” *Management Science*, 2015, 62 (3), 631–647.
- Phillips, Lawrence D and Ward Edwards**, “Conservatism in a simple probability inference task,” *Journal of experimental psychology*, 1966, 72 (3), 346.
- Piff, Paul K, Daniel M Stancato, Stéphane Côté, Rodolfo Mendoza-Denton, and Dacher Keltner**, “Higher social class predicts increased unethical behavior,” *Proceedings of the National Academy of Sciences*, 2012, 109 (11), 4086–4091.

- Qiao, Baoyun, Jorge Martinez-Vazquez, and Yongsheng Xu**, “The tradeoff between growth and equity in decentralization policy: China’s experience,” *Journal of Development Economics*, 2008, 86 (1), 112–128.
- Rao, Gautam**, “Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools,” *American Economic Review*, 2019, 109 (3), 774–809.
- Rohner, Dominic, Mathias Thoenig, and Fabrizio Zilibotti**, “Seeds of distrust: Conflict in Uganda,” *Journal of Economic Growth*, 2013, 18 (3), 217–252.
- Ross, Lee**, “The intuitive psychologist and his shortcomings: Distortions in the attribution process,” *Advances in experimental social psychology*, 1977, 10, 173–220.
- Rubin, Donald B**, “Randomization analysis of experimental data: The Fisher randomization test comment,” *Journal of the American Statistical Association*, 1980, 75 (371), 591–593.
- Scharping, Thomas**, *Birth control in China 1949-2000: population policy and demographic development*, Routledge, July 2013.
- Serrato, Juan Carlos Suárez, Xiao Yu Wang, and Shuang Zhang**, “The one child policy and promotion of mayors in China,” *NBER Working Paper*, 2016.
- Shirk, Susan L**, *Changing media, changing China*, Oxford University Press, 2011.
- Storeygard, Adam**, “Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa,” *The Review of Economic Studies*, 2016, 83 (3), 1263–1295.
- Tang, Wei and Geoffrey JD Hewings**, “Do city–county mergers in China promote local economic development?,” *Economics of Transition*, 2017, 25 (3), 439–469.
- Vaubel, Roland**, “A Public choice approach to international organization,” *Public Choice*, 1986, 51 (1), 39–57.

- Wang, Long and J Keith Murnighan**, “Money, emotions, and ethics across individuals and countries,” *Journal of Business Ethics*, 2014, 125 (1), 163–176.
- Wang, Zhen**, “Who gets promoted and why? Understanding power and persuasion in China’s cadre evaluation system,” in “Annual Meeting of the American Association for Chinese Studies, New Brunswick, New Jersey” 2013.
- Watson, Andrew, Christopher Findlay, and Du Yintang**, “Who won the “ wool war”?: A case study of rural product marketing in China,” *China Quarterly*, 1989, pp. 213–241.
- Wei, Shang-Jin and Xiaobo Zhang**, “The competitive saving motive: Evidence from rising sex ratios and savings rates in China,” *Journal of Political Economy*, 2011, 119 (3), 511–564.
- **and** – , “Sex ratios, entrepreneurship, and economic growth in the People’s Republic of China,” *NBER Working Paper 16800*, 2011.
- Wilson, Kenneth and Ian CW Hardy**, “Statistical analysis of sex ratios: An introduction,” *Sex Ratios: Concepts and Research Methods*, 2002, 1, 48–92.
- Wolfers, Justin**, *Are voters rational? Evidence from gubernatorial elections*, Graduate School of Business, Stanford University, 2002.
- Xie, Danyang, Heng fu Zou, and Hamid Davoodi**, “Fiscal decentralization and economic growth in the United States,” *Journal of Urban economics*, 1999, 45 (2), 228–239.
- Xu, Chenggang**, “The fundamental institutions of China’s reforms and development,” *Journal of Economic Literature*, 2011, 49 (4), 1076–1151.
- Young, Alwyn**, “The razor’s edge: Distortions and incremental reform in the People’s Republic of China,” *The Quarterly Journal of Economics*, 2000, 115 (4), 1091–1135.
- Zhang, Tao and Heng fu Zou**, “Fiscal decentralization, public spending, and economic growth in China,” *Journal of Public Economics*, 1998, 67 (2), 221–240.

Zhao, Yuezhi, *Media, market, and democracy in China: Between the Party line and the bottom line*, Vol. 137, University of Illinois Press, 1998.